Soft-In Soft-Out DFE and Bi-directional DFE

Seongwook Jeong, *Student Member, IEEE* and Jaekyun Moon[†], *Fellow, IEEE*

Dept. of Electrical and Computer Engineering

University of Minnesota

Minneapolis, Minnesota 55455, U.S.A.

Email: jeong030@umn.edu

[†] Dept. of Electrical Engineering

Korea Advanced Institute of Science and Technology

Daejeon, 305-701, Republic of Korea

Email: jmoon@kaist.edu

Abstract

We design a soft-in soft-out (SISO) decision feedback equalizer (DFE) that performs better than its linear counterpart in turbo equalizer (TE) setting. Unlike previously developed SISO-DFEs, the present DFE scheme relies on extrinsic information formulation that directly takes into account the error propagation effect. With this new approach, both error rate simulation and the extrinsic information transfer (EXIT) chart analysis indicate that the proposed SISO-DFE is superior to the well-known SISO linear equalizer (LE). This result is in contrast with the general understanding today that the error propagation effect of the DFE degrades the overall TE performance below that of the TE based on a LE. We also describe a new extrinsic information combining strategy involving the outputs of two DFEs running in opposite directions, that explores error correlation between the two sets of DFE outputs. When this method is combined with the new DFE extrinsic information formulation, the resulting "bidirectional" turbo-DFE achieves excellent performance-complexity tradeoffs compared to the TE based on the BCJR algorithm or on the LE. Unlike turbo LE or turbo DFE, the turbo BiDFE's performance does not degrade significantly as the feedforward and feedback filter taps are constrained to be time-invariant.

I. INTRODUCTION

Intersymbol interference (ISI) arises as the transmitted symbols overlaps with one another in high speed digital communication. Powerful modern equalization methods are based on the turbo equalization principle established in [1], wherein a soft-in soft-out (SISO) equalizer (or detector) and a SISO error-correction decoder exchange soft information in an iterative fashion until reliable decisions are generated. It has been shown in [1] that even for some heavy ISI channels the detrimental effect of ISI disappears with this approach.

The detector or the equalizer portion of a turbo equalizer (TE) system often investigated is based on the well-known Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [2]. This algorithm exactly computes the *a posteriori* probability (APP) of the transmitted signal symbols considering the channel response and the *a priori* information of the transmitted symbols and, as such, can be viewed as an optimum SISO equalizer. However, the computational complexity of this algorithm grows exponentially as a function of the channel length and the symbol alphabet set size.

The high computational complexity of the BCJR-based equalizer has motivated considerable research on numerous suboptimal but low complexity equalization schemes. A notable development along this direction is the well-known SISO linear equalizer (LE) of [3]. Another possibility, which was also evaluated in [3], is the SISO decision feedback equalizer (DFE). In the classical, non-turbo setting (i.e., no iterative exchange of soft information between the equalizer and the decoder), it has long been known that the DFE almost always outperforms the LE, despite the fact that the DFE typically suffers from error propagation. This is because when ISI is severe with the channel response showing nulls or deep valleys within the Nyquist band, the LE is subject to large noise enhancement. The work of [3], however, shows that when hard decisions are fed through the feedback filter (to reduce complexity), SISO-DFE performs considerably worse than SISO-LE, presumably due to error propagation.

In classical DFE setting, many techniques have been investigated to mitigate error propagation [4], [5], [6]. Recently, it has been shown [7], [8], [9] that conducting both normal and time-reversed equalization of the received data sequence with two DFEs running in opposite directions and combining two DFE outputs is very effective in reducing error propagation and improving bit error rate (BER) performance. This "bi-directional" DFE (called BiDFE) algorithm takes advantage of the different decision error and noise distributions at the outputs of the forward and time-reversed DFEs [7], [8].

The contribution of this paper is two-fold. One is that this paper readdresses the DFE design issue in the turbo equalizer environment and shows that just as in classical non-turbo setting, the DFE outperforms

the LE, if extrinsic information is reformulated in a way that combats error propagation more effectively. The second contribution is a specific DFE extrinsic information combining strategy applied to a BiDFE that suppresses statistical correlation between the outputs of two opposite direction DFEs. We show that the resulting turbo BiDFE performance approaches the performance of the BCJR-based turbo equalizer in a fairly severe ISI environment, easily outperforming the turbo equalizer based on the SISO-LE of [3]. Remarkably, the performance of a time-invariant version of the BiDFE, a lower-complexity method that does not require tap-weight updating as a function of time, also consistently is better than the SISO-LE scheme of [3] based on a time-varying linear filter. There also exist feedback equalization techniques that utilize soft decisions to reduce error propagation [6], [9], [10], [11] but we focus on hard-decision feedback in this paper, as the feedback finite-impulse-response filter complexity is greatly reduced when feedback decisions are constrained to take hard values.

The remainder of the paper is organized as follows. In Section II, a brief statement of the problem is given. In Section III, we give a quick review of the SISO equalizer design method established in [3] and then provide a new formulation of the extrinsic information of DFE taking into account the error propagation effect. We also provide the mean-squared-error analysis of the infinite-length BiDFE in Section IV. The iterative BiDFE algorithm is introduced with the extrinsic information combiner of the normal forward and time-reversed DFE outputs in Section V. In Section VI, numerical results and analysis are given. Finally, we draw conclusions in Section VII.

II. SYSTEM MODEL

We assume that the receiver knows the discrete-time baseband channel response accurately. While the methods discussed are general, our presentation will be based on binary symbols with $P_x \triangleq \mathrm{E}(x_n^2) = 1$, $x_n \in \{\pm 1\}$, as well as real-valued ISI channel coefficients and noise samples. Although x_n typically represents a coded bit sequence, our analysis will assume that it is equiprobable and independent and identically distributed (i.i.d.). Given the transmitted bit sequence $\{x_k\}$, the channel output at time n is

$$r_n = \sum_{k=0}^{L_h-1} h_k x_{n-k} + w_n \tag{1}$$

where w_n is additive white Gaussian noise (AWGN) with variance N_0 and $\{h_k\}$ is the channel impulse response with length L_h.

In turbo equalization, the equalizer computes the *a posteriori* log-likelihood ratio (LLR) of x_n,

$$L(x_n) \triangleq \ln \frac{\Pr(x_n = +1 \mid \mathbf{r}_n)}{\Pr(x_n = -1 \mid \mathbf{r}_n)}$$

where $\mathbf{r}_n$ is the received sample block utilized for LLR estimation for x_n. Note that this computation requires the knowledge of the *a priori* probabilities of all input bits affecting $\mathbf{r}_n$. Since these *a priori* probabilities are not available, they are all set to 1/2 initially and then, as the turbo iteration ensues, to the estimated probability values based on the extrinsic information generated and passed back by the outer decoder.

The equalizer then generates its own extrinsic information by subtracting the effect of the probability estimate passed down for the current bit. Write this estimated *a priori* LLR passed down from the decoder as

$$L_a(x_n) \triangleq \ln \frac{\Pr(x_n = +1)}{\Pr(x_n = -1)}$$

with an understanding that the probabilities in the expression are in reality just estimates.

Then, the equalizer's extrinsic LLR for x_n to be passed to the error-correction code decoder is given by

$$L_e(x_n) \triangleq L(x_n) - L_a(x_n).$$

This equation suggests first computing $L(x_n)$ based on the a priori probabilities of all input bits including x_n and then simply subtracting $L_a(x_n)$ to generate the extrinsic LLR $L_e(x_n)$. An alternative way of generating $L_e(x_n)$ is to set $L_a(x_n) = 0$ while computing $L(x_n)$, i.e., suppress the effect of $L_a(x_n)$ in the calculation of $L(x_n)$:

$$L_e(x_n) = L(x_n)|_{L_a(x_n)=0}.$$

The techniques discussed in this paper actually use the second method.

III. DERIVATION OF MODIFIED ITERATIVE DFE ALGORITHM

In this section we first briefly review the results of [3] related to the SISO-DFE to provide necessary background while establishing notation. We then show a new way of computing extrinsic information so as to suppress error propagation and improve performance.

A. Review of Existing Extrinsic LLR Mapping

The work of [3] has established an effective strategy of utilizing the *a priori* information estimates from the outer decoder in calculating the equalizer tap coefficients. The gist of the approach in [3] is a clever tweaking of the classical minimum-mean-squared-error (MMSE) estimation principle where the "mean" of the input symbols are constructed using the available *a priori* information estimates and

utilized in the linear estimator weight computation. Both the LE and the DFE can be designed in this way, but we shall focus on the DFE here. Based on the above principle and suppressing the effect of the *a priori* probability estimate on the current bit x_n (i.e., $\mathrm{E}(x_n) = 0$) in an effort to extract the extrinsic information, the MMSE feedforward filter taps (a total of $L_c + 1$) and the feedback filter taps (a total of $L_d = L_h - 1$) at time n are derived respectively as:

$$\mathbf{c}_n \triangleq \left[c_{\{n,0\}}, c_{\{n,+1\}}, \ldots, c_{\{n,L_c\}} \right]^T$$

$$= \left\{ \mathbf{H}\mathbf{\Sigma}_n\mathbf{H}^T + (1 - z_n)\mathbf{s}\mathbf{s}^T + N_0\mathbf{I} \right\}^{-1} \mathbf{s} \tag{2}$$

$$\mathbf{d}_n \triangleq \left[d_{\{n,-L_d\}}, d_{\{n,-L_d+1\}}, \ldots, d_{\{n,-1\}} \right]^T$$

$$= \mathbf{M}\mathbf{H}^T\mathbf{c}_n \tag{3}$$

where $\mathbf{H}$ is a channel convolution matrix defined as

$$\mathbf{H} \triangleq \begin{bmatrix} h_{L_h-1} & h_{L_h-2} & \cdots & h_0 & 0 & \cdots & & 0 \\ 0 & h_{L_h-1} & h_{L_h-2} & \cdots & h_0 & 0 & \cdots & 0 \\ & \ddots & & & \ddots & & \ddots & \\ 0 & 0 & \cdots & 0 & h_{L_h-1} & h_{L_h-2} & \cdots & h_0 \end{bmatrix}$$

and the matrix $\mathbf{\Sigma}_n$ depends on $\mathrm{E}(x_i)$, $i = n, n + 1, \ldots, n + L_c$, computed from the decoder output as $\mathrm{E}(x_i) = \tanh(L_a(x_i)/2)$. Specifically, $\mathbf{\Sigma}_n \triangleq \mathrm{Diag}(\mathbf{0}_{1 \times L_d}, z_n, z_{n+1}, \ldots, z_{n+L_c})$ with $z_i \triangleq 1 - [\mathrm{E}(x_i)]^2$. Adding the term $(1 - z_n)\mathbf{s}\mathbf{s}^T$ in (2) has the same effect of suppressing $\mathrm{E}(x_n)$ to zero in $\mathbf{H}\mathbf{\Sigma}_n\mathbf{H}^T$. The remaining vector and matrix are defined as $\mathbf{s} \triangleq \mathbf{H}[\mathbf{0}_{1 \times L_d}, 1, \mathbf{0}_{1 \times L_c}]^T$ and $\mathbf{M} \triangleq [\mathbf{I}_{L_d \times L_d}, \mathbf{0}_{L_d \times (L_c+1)}]$.

The equalizer output is obtained as

$$y_n = \mathbf{c}_n^T \cdot (\mathbf{r}_n - \mathbf{H}\bar{\mathbf{x}}_n + \mathrm{E}(x_n)\mathbf{s}) \tag{4}$$

where the received vector is defined as $\mathbf{r}_n \triangleq [r_n, r_{n+1}, \ldots, r_{n+L_c}]^T$ and the composite vector of the causal symbol decisions and the anticausal symbols' mean as $\bar{\mathbf{x}}_n \triangleq [\hat{x}_{n-L_d}, \ldots, \hat{x}_{n-1}, \mathrm{E}(x_n), \ldots, \mathrm{E}(x_{n+L_c})]^T$ where $\hat{x}_i$ is the available decision for x_i based on the *a posteriori* LLR of x_i, i.e., if $L(x_i) = L_a(x_i) + L_e(x_i) \geq 0$, then, $\hat{x}_i = +1$; otherwise, $\hat{x}_i = -1$. The addition of the $\mathrm{E}(x_n)\mathbf{s}$ term is also to suppress the effect of $\mathrm{E}(x_n)$ in $\mathbf{H}\bar{\mathbf{x}}_n$.

Define the anticausal symbol sequence $\mathbf{x}_n \triangleq [x_n, x_{n+1}, \ldots, x_{n+L_c}]^T$, the causal symbol sequence $\mathbf{x}_n^c \triangleq [x_{n-L_d}, x_{n-L_d+1}, \ldots, x_{n-1}]^T$, and the available decision sequence $\hat{\mathbf{x}}_n^c \triangleq [\hat{x}_{n-L_d}, \hat{x}_{n-L_d+1}, \ldots, \hat{x}_{n-1}]^T$. Also define the noise sequence as $\mathbf{w}_n \triangleq [w_n, w_{n+1}, \ldots, w_{n+L_c}]^T$. Then, the combined filter output y_n

can be rewritten as

$$y_n = (\mathbf{c}_n^T \mathbf{H}_1) \cdot \left(\mathbf{x}_n - \mathrm{E}\{\dot{\mathbf{x}}_n\} \right) + \mathbf{d}_n^T (\mathbf{x}_n^c - \hat{\mathbf{x}}_n^c) + \mathbf{c}_n^T \mathbf{w}_n$$

$$= p_{\{n,0\}} x_n + \sum_{k=1}^{L_d} d_{\{n,-k\}} \left(x_{n-k} - \hat{x}_{n-k} \right) + \sum_{k=1}^{L_c} p_{\{n,k\}} \left(x_{n+k} - \mathrm{E}(x_{n+k}) \right) + \sum_{k=0}^{L_c} c_{\{n,k\}} w_{n+k}$$

$$= p_{\{n,0\}} x_n + i_n + v_n \tag{5}$$

where $\mathrm{E}\{\dot{\mathbf{x}}_n\} \triangleq [0, \mathrm{E}(x_{n+1}), \mathrm{E}(x_{n+2}), \dots, \mathrm{E}(x_{n+L_c})]^T$ and $\mathbf{H}_1$ is the $(L_c + 1) \times (L_c + 1)$ submatrix of $\mathbf{H}$ formed by the entire rows of the columns from the $(L_d + 1)$th to the last. Moreover, $\mathbf{p}_n \triangleq \left[p_{\{n,0\}}, p_{\{n,1\}}, \dots, p_{\{n,L_c\}} \right] = \mathbf{c}_n^T \mathbf{H}_1$ and $p_{\{n,0\}} = \mathbf{c}_n^T \mathbf{s}$. The error propagation caused by the mismatched hard decision feedback is denoted as i_n, i.e., $i_n \triangleq \sum_{k=1}^{L_d} d_{\{n,-k\}} \left(x_{n-k} - \hat{x}_{n-k} \right)$ and v_n is the sum of noise and the remaining ISI terms caused by the neighboring symbols: $v_n \triangleq \sum_{k=1}^{L_c} p_{\{n,k\}} \left(x_{n+k} - \mathrm{E}(x_{n+k}) \right) + \sum_{k=0}^{L_c} c_{\{n,k\}} w_{n+k}$. The variance of v_n is

$$\mathrm{Var}(v_n) \triangleq \mathbf{c}_n^T \mathrm{Cov}\{\mathbf{r}_n \mathbf{r}_n^T \mid x_n = x\} \mathbf{c}_n$$

$$= \mathbf{c}_n^T \mathbf{s}(1 - \mathbf{s}^T \mathbf{c}_n). \tag{6}$$

Assuming that the feedback decisions are all correct, i.e., $i_n = 0$, and v_n is AWGN, the extrinsic LLR is naturally given by

$$L_e(x_n) \triangleq \ln \frac{\Pr(x_n = +1 \mid y_n)}{\Pr(x_n = -1 \mid y_n)} \bigg|_{L_a(x_n)=0}$$

$$= \ln \frac{\Pr(y_n \mid x_n = +1)\Pr(x_n = +1)}{\Pr(y_n \mid x_n = -1)\Pr(x_n = -1)} \bigg|_{L_a(x_n)=0}$$

$$= \ln \frac{\Pr(y_n \mid x_n = +1)}{\Pr(y_n \mid x_n = -1)}$$

$$= -\frac{\left| y_n - p_{\{n,0\}} \right|^2}{2\mathrm{Var}(v_n)} + \frac{\left| y_n + p_{\{n,0\}} \right|^2}{2\mathrm{Var}(v_n)}$$

$$= \frac{2 p_{\{n,0\}} y_n}{\mathrm{Var}(v_n)}. \tag{7}$$

Notice that in generating y_n, $L_a(x_n)$ was already suppressed to zero.

A glossary of frequently used symbols is given below. Time-varying quantities are augmented with time index n as the subscript.

B. New Formulation of Extrinsic Information

While the MAP estimation of i_n is equal to zero, we observe that the chance of $i_n \neq 0$ is relatively high for severe ISI channels. Our strategy is to estimate i_n and utilize the statistical parameters associated

$\mathbf{c}_n$	DFE feedforward filter coefficients of length $L_c + 1$	x_n	transmitted symbol
$\mathbf{d}_n$	DFE feedback filter coefficients of length L_d	w_n	channel noise
$\mathbf{H}$	channel convolution matrix	P_x	average power of x_n
$\mathbf{M}$	$[\mathbf{I}_{L_d \times L_d}, \mathbf{0}_{L_d \times (L_c+1)}]$	N_0	variance of w_n
$\mathbf{s}$	$\mathbf{H}[\mathbf{0}_{1 \times L_d}, 1, \mathbf{0}_{1 \times L_c}]^T$	$\{h_k\}$	ISI channel response of length L_h
$\mathbf{p}_n$	$\mathbf{c}_n^T \mathbf{H}_1$ where $\mathbf{H}_1$ is a submatrix of $\mathbf{H}$	r_n	received channel output
$\mathbf{r}_n$	received sample vector	y_n	equalized observation
$\bar{\mathbf{x}}_n$	vector of causal decisions and anticausal's mean	i_n	error due to mismatched past decisions
$\mathbf{w}_n$	noise sample vector	v_n	noise plus error due to pre-cursor ISI
$\mathbf{x}_n$	transmitted anticausal symbol vector	$p_{\{n,0\}}$	weight on x_n in y_n
$\mathbf{x}_n^c$	transmitted causal symbol vector	$L_a(x_n)$	*a priori* LLR of x_n
$\hat{\mathbf{x}}_n^c$	estimated causal symbol vector	$L(x_n)$	*a posteriori* LLR of x_n
$\mathbf{y}_n^c$	equalized causal sample vector	$L_e(x_n)$	extrinsic LLR of x_n
$\mathbf{e}_{\{n,j\}}^c$	possible causal error sequence	z_n	variance of x_n
$\mathbf{\Sigma}_n$	covariance matrix of transmitted anticausal symbols	$\acute{z}_n$	variance of x_n estimated via *a posteriori* LLR
$\acute{\mathbf{\Sigma}}_n^c$	covariance matrix of estimated causal symbols	ρ_n	noise correlation coefficient between two DFEs

with this estimate in the formulation of the extrinsic information. Since i_n is to be estimated on the basis of the observation $\mathbf{y}_n^c \triangleq [y_{n-L_d}, y_{n-L_d+1}, \ldots, y_{n-1}]^T$, the mean and variance of i_n can be evaluated by the *a posteriori* probabilities of the causal symbols. Write

$$\mathrm{E}(i_n) \triangleq \mathrm{E}\left\{\mathbf{d}_n^T(\mathbf{x}_n^c - \hat{\mathbf{x}}_n^c) \mid \mathbf{y}_n^c\right\}$$

$$= \mathbf{d}_n^T\left(\tanh(L(\mathbf{x}_n^c)/2) - \hat{\mathbf{x}}_n^c\right) \tag{8}$$

$$\mathrm{Var}(i_n) \triangleq \mathrm{Var}\left\{\mathbf{d}_n^T(\mathbf{x}_n^c - \hat{\mathbf{x}}_n^c) \mid \mathbf{y}_n^c\right\}$$

$$= \mathbf{d}_n^T \acute{\mathbf{\Sigma}}_n^c \mathbf{d}_n \tag{9}$$

where $L(\mathbf{x}_n^c) = [L(x_{n-L_d}), L(x_{n-L_d+1}), \ldots, L(x_{n-1})]^T$, $\acute{\mathbf{\Sigma}}_n^c \triangleq \mathrm{Diag}\left(\acute{z}_{n-L_d}, \acute{z}_{n-L_d+1}, \ldots, \acute{z}_{n-1}\right)$, and $\acute{z}_n = 1 - \tanh(L(x_n)/2)^2$.

Now, let us consider the possible causal error sequence $\mathbf{e}_{\{n,j\}}^c \triangleq \mathbf{x}_{\{n,j\}}^c - \hat{\mathbf{x}}_n^c$ for $j = 1, 2, \ldots, 2^{L_d}$, with index j pointing to a particular binary pattern of $\mathbf{x}_n^c$. Then, we can compute the extrinsic information for the given causal error sequence $\mathbf{e}_{\{n,j\}}^c$:

$$L_e(x_n|\mathbf{e}_{\{n,j\}}^c) \triangleq \ln \frac{\Pr(y_n \mid x_n = +1, \mathbf{e}_{\{n,j\}}^c)}{\Pr(y_n \mid x_n = -1, \mathbf{e}_{\{n,j\}}^c)}$$

$$= \frac{2p_{\{n,0\}}(y_n - \mathbf{d}_n^T \mathbf{e}_{\{n,j\}}^c)}{\mathrm{Var}(v_n)}. \tag{10}$$

To compute the extrinsic information of x_n taking into account the probabilities of possible error sequences, we write

$$
\begin{aligned}
\Pr(y_n \mid x_n = +1) &= \sum_{j=1}^{2^{L_d}} \Pr(y_n \mid x_n = +1, \mathbf{e}^c_{\{n,j\}}) \Pr(\mathbf{e}^c_{\{n,j\}}) \\
&= \sum_{j=1}^{2^{L_d}} \frac{\exp\left(L_e(x_n|\mathbf{e}^c_{\{n,j\}})\right) \Pr(\mathbf{e}^c_{\{n,j\}})}{1 + \exp\left(L_e(x_n|\mathbf{e}^c_{\{n,j\}})\right)}
\end{aligned} \tag{11}
$$

$$
\begin{aligned}
\Pr(y_n \mid x_n = -1) &= \sum_{j=1}^{2^{L_d}} \Pr(y_n \mid x_n = -1, \mathbf{e}^c_{\{n,j\}}) \Pr(\mathbf{e}^c_{\{n,j\}}) \\
&= \sum_{j=1}^{2^{L_d}} \frac{\Pr(\mathbf{e}^c_{\{n,j\}})}{1 + \exp\left(L_e(x_n|\mathbf{e}^c_{\{n,j\}})\right)}.
\end{aligned} \tag{12}
$$

Accordingly, the extrinsic information of x_n considering the distribution of i_n is given as

$$
L_e(x_n) = \ln\left\{ \sum_{j=1}^{2^{L_d}} \frac{\exp\left(L_e(x_n|\mathbf{e}^c_{\{n,j\}})\right) \Pr(\mathbf{e}^c_{\{n,j\}})}{1 + \exp\left(L_e(x_n|\mathbf{e}^c_{\{n,j\}})\right)} \right\} - \ln\left\{ \sum_{j=1}^{2^{L_d}} \frac{\Pr(\mathbf{e}^c_{\{n,j\}})}{1 + \exp\left(L_e(x_n|\mathbf{e}^c_{\{n,j\}})\right)} \right\}. \tag{13}
$$

In principle, the extrinsic information of (13) can be evaluated using (10) and approximating $\Pr(\mathbf{e}^c_{\{n,j\}})$ or $\Pr(\mathbf{e}^c_{\{n,j\}}|\mathbf{y}^c_n)$ by $\prod_{k=1}^{L_d} \Pr(e_{\{n-k,j\}}|y_{n-k})$, which can be computed based on the *a posteriori* LLRs of $\mathbf{x}^c_n$.

However, since the computational complexity of (13) increases exponentially according to the length of feedback filter, L_d, we seek a more practical modification. A possible solution is to apply the Bayes' rule only for the two mutually exclusive cases of $i_n = 0$ and $i_n \neq 0$. Then,

$$
\Pr(y_n \mid x_n = +1) = \frac{\exp\left(L_e(x_n|i_n = 0)\right) \Pr(i_n = 0)}{1 + \exp\left(L_e(x_n|i_n = 0)\right)} + \frac{\exp\left(L_e(x_n|i_n \neq 0)\right) \Pr(i_n \neq 0)}{1 + \exp\left(L_e(x_n|i_n \neq 0)\right)} \tag{14}
$$

$$
\Pr(y_n \mid x_n = -1) = \frac{\Pr(i_n = 0)}{1 + \exp\left(L_e(x_n|i_n = 0)\right)} + \frac{\Pr(i_n \neq 0)}{1 + \exp\left(L_e(x_n|i_n \neq 0)\right)}. \tag{15}
$$

The extrinsic information of x_n for each case of i_n can be estimated as

$$L_e(x_n|i_n = 0) = \frac{2p_{\{n,0\}}y_n}{\text{Var}(v_n)} \tag{16}$$

$$L_e(x_n|i_n \neq 0) = \ln\left\{\sum_{j=1,\mathbf{e}_{\{n,j\}}^c \neq \mathbf{0}}^{2^{L_d}} \frac{\exp\left(L_e(x_n|\mathbf{e}_{\{n,j\}}^c)\right)\text{Pr}(\mathbf{e}_{\{n,j\}}^c)}{\left\{1 + \exp\left(L_e(x_n|\mathbf{e}_{\{n,j\}}^c)\right)\right\}\text{Pr}(i_n \neq 0)}\right\}$$

$$- \ln\left\{\sum_{j=1,\mathbf{e}_{\{n,j\}}^c \neq \mathbf{0}}^{2^{L_d}} \frac{\text{Pr}(\mathbf{e}_{\{n,j\}}^c)}{\left\{1 + \exp\left(L_e(x_n|\mathbf{e}_{\{n,j\}}^c)\right)\right\}\text{Pr}(i_n \neq 0)}\right\}$$

$$\simeq \ln\left\{\sum_{j=1,\mathbf{e}_{\{n,j\}}^c \neq \mathbf{0}}^{2^{L_d}} \left(\frac{1}{2} + \frac{L_e(x_n|\mathbf{e}_{\{n,j\}}^c)}{4}\right)\frac{\text{Pr}(\mathbf{e}_{\{n,j\}}^c)}{\text{Pr}(i_n \neq 0)}\right\}$$

$$- \ln\left\{\sum_{j=1,\mathbf{e}_{\{n,j\}}^c \neq \mathbf{0}}^{2^{L_d}} \left(\frac{1}{2} - \frac{L_e(x_n|\mathbf{e}_{\{n,j\}}^c)}{4}\right)\frac{\text{Pr}(\mathbf{e}_{\{n,j\}}^c)}{\text{Pr}(i_n \neq 0)}\right\} \tag{17}$$

$$= \ln\left\{\mathop{\text{E}}_{i_n}\left(\frac{1}{2} + \frac{2p_{\{n,0\}}(y_n - i_n)}{4\text{Var}(v_n)}\bigg| i_n \neq 0\right)\right\} - \ln\left\{\mathop{\text{E}}_{i_n}\left(\frac{1}{2} - \frac{2p_{\{n,0\}}(y_n - i_n)}{4\text{Var}(v_n)}\bigg| i_n \neq 0\right)\right\}$$

$$= \ln\left\{1 + \frac{p_{\{n,0\}}(y_n - \text{E}(i_n|i_n \neq 0))}{\text{Var}(v_n)}\right\} - \ln\left\{1 - \frac{p_{\{n,0\}}(y_n - \text{E}(i_n|i_n \neq 0))}{\text{Var}(v_n)}\right\}$$

$$\simeq \begin{cases} 2\varphi_n/(1 - \varphi_n) \text{ if } \varphi_n < 0 \\ 2\varphi_n/(1 + \varphi_n) \text{ otherwise} \end{cases} \tag{18}$$

$$= \frac{2\varphi_n}{1 + |\varphi_n|} \tag{19}$$

where $\varphi_n \triangleq p_{\{n,0\}}(y_n - \text{E}(i_n|i_n \neq 0))/\text{Var}(v_n)$, $\text{E}(i_n|i_n \neq 0) = \text{E}(i_n)/\text{Pr}(i_n \neq 0)$, $\text{Pr}(i_n = 0) = \prod_{k=1}^{L_d}\exp(|L(x_{n-k})|)/(1 + \exp(|L(x_{n-k})|))$, and $\text{Pr}(i_n \neq 0) = 1 - \text{Pr}(i_n = 0)$. The approximation of (17) is from the first order Taylor expansion at zero, i.e, $e^x/(1 + e^x) \simeq 0.5 + 0.25x$ and $1/(1 + e^x) \simeq 0.5 - 0.25x$. Furthermore, we also use $\ln\{1 + \varphi_n\} - \ln\{1 - \varphi_n\} = \ln\{1 + 2\varphi_n/(1 - \varphi_n)\} = -\ln\{1 - 2\varphi_n/(1 + \varphi_n)\}$ and $\ln(1+x) \simeq x$ in (18). In other words, $\ln\{1 + 2\varphi_n/(1 - \varphi_n)\} \simeq 2\varphi_n/(1 - \varphi_n)$ is used for $\varphi_n < 0$ while $-\ln\{1 - 2\varphi_n/(1 + \varphi_n)\} \simeq 2\varphi_n/(1 + \varphi_n)$ is used for $\varphi_n \geq 0$.

Finally, the extrinsic information of x_n is given as

$$L_e(x_n) = \ln\left\{\frac{\exp(L_e(x_n|i_n = 0))\text{Pr}(i_n = 0)}{1 + \exp(L_e(x_n|i_n = 0))} + \frac{\exp(L_e(x_n|i_n \neq 0))\text{Pr}(i_n \neq 0)}{1 + \exp(L_e(x_n|i_n \neq 0))}\right\}$$

$$- \ln\left\{\frac{\text{Pr}(i_n = 0)}{1 + \exp(L_e(x_n|i_n = 0))} + \frac{\text{Pr}(i_n \neq 0)}{1 + \exp(L_e(x_n|i_n \neq 0))}\right\}. \tag{20}$$

While this gets passed to the outer decoder as equalizer's extrinsic information, hard decisions that propagate down the feedback filter are generated by slicing $L_e(x_n) + L_a(x_n)$ where $L_a(x_n)$ is the extrinsic information from the decoder.

C. Time-Invariant Filters

As also discussed in [3], the filter tap values derived above are time-varying and creates significant implementation challenges. A low-complexity variation would be to simply assume the classical (non-turbo) DFE forward and feedback filter tap solutions as in

$$\mathbf{c} \triangleq [c_0, c_{+1}, \ldots, c_{L_c}]^T$$

$$= \left(\mathbf{H\Sigma H}^T + N_0\mathbf{I}\right)^{-1}\mathbf{s} \tag{21}$$

$$\mathbf{d} \triangleq [d_{-L_d}, d_{-L_d+1}, \ldots, d_{-1}]^T$$

$$= \mathbf{MH}^T\mathbf{c}, \tag{22}$$

where $\mathbf{\Sigma} \triangleq \mathrm{Diag}(\mathbf{0}_{1\times L_d}, \mathbf{1}_{1\times(L_c+1)})$, but let the effect of decoder feedback come into play through the subtraction of $\mathbf{H\bar{x}}_n - \mathrm{E}(x_n)\mathbf{s}$ from the channel observation vector (see (4)) and the enhanced *a posteriori* LLR computation: $L_e(x_n) + L_a(x_n)$ where $L_a(x_n)$ represents the decoder feedback.

By an obvious modification of (5), the equalized signal is obtained as

$$y_n = p_0 x_n + i_n + v_n \tag{23}$$

where $p_0 = \mathbf{c}^T\mathbf{s}$, $i_n = \sum_{k=1}^{L_d} d_{-k}\left(x_{n-k} - \hat{x}_{n-k}\right)$, $v_n = \sum_{k=1}^{L_c} p_k(x_{n+k} - \mathrm{E}(x_{n+k})) + \sum_{k=0}^{L_c} c_k w_{n+k}$, and $\mathbf{p} \triangleq [p_0, p_1, \ldots, p_{L_c}] = \mathbf{c}^T\mathbf{H}_1$. The mean and variance of i_n and the noise variance of v_n with the time-invariant filters are also given by

$$\mathrm{E}(i_n) = \mathbf{d}^T\left(\tanh(L(\mathbf{x}_n^c)/2) - \hat{\mathbf{x}}_n^c\right) \tag{24}$$

$$\mathrm{Var}(i_n) = \mathbf{d}^T\acute{\mathbf{\Sigma}}_n^c\mathbf{d} \tag{25}$$

$$\mathrm{Var}(v_n) = \mathbf{c}^T\left(\mathbf{H\Sigma}_n\mathbf{H}^T - z_n\mathbf{ss}^T + N_0\mathbf{I}\right)\mathbf{c}. \tag{26}$$

IV. SNR ADVANTAGE OF BiDFE

The idea of BiDFE is already motivated in [7], [8] by the fact that DFE can be performed on the reversed received sequence using the time-reversed channel response. Here we derive the SNR figure-of-merit for BiDFE assuming ideal feedback in both ways and allowing infinitely long filter lengths. We then compare the result with those of the usual, single-sided DFE as well as the matched filter detector (i.e., ideal detector under zero-ISI condition). As will be seen, the ideal BiDFE SNR is significantly better than the ideal DFE SNR especially at high channel SNRs, further motivating a turbo BiDFE scheme.

A. Unbiased MMSE-DFE

It is well known that the D-transforms of the feedforward and feedback MMSE-DFE filter coefficients are, respectively [12]:

$$c(D) = \frac{P_x}{P_0 g^*(D^{-*})}, \qquad d(D) = g(D) \tag{27}$$

where P_0 is such that $\log P_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log R_{ss}(e^{-j\theta}) d\theta$ and $g^*(D^{-*})$ is obtained from spectral factorization: $R_{ss}(D) = P_x R_{hh}(D) + N_0 = P_0 g(D) g^*(D^{-*})$ where $R_{hh}(D) = h(D) h^*(D^{-*})$ and $h(D)$ is the D-transform of the channel impulse response.

The unbiased equalized outputs of the normal MMSE-DFE in the forward direction, $Y_f(D)$, are given by

$$Y_f(D) = x(D) + \frac{P_0}{P_0 - N_0} e'_f(D) \tag{28}$$

where

$$e'_f(D) \triangleq \frac{N_0}{P_0} \left(1 - \frac{1}{g^*(D^{-*})}\right) x(D) + \frac{P_x w'(D)}{P_0 g^*(D^{-*})} \tag{29}$$

with $w'(D)$ denoting a complex-valued Gaussian noise sequence with autocorrelation function $R_{w'w'}(D) = N_0 R_{hh}(D)$. Then, the mean-squared-error (MSE) and SNR of the unbiased normal MMSE-DFE are given by

$$\text{MSE}_{UDFE} = \left(\frac{P_0}{P_0 - N_0}\right)^2 \text{E}(|e'_{f,n}|^2) = \frac{P_x N_0}{P_0 - N_0} \tag{30}$$

$$\text{SNR}_{UDFE} \triangleq \frac{P_x}{\text{MSE}_{UDFE}} = \frac{P_0 - N_0}{N_0}. \tag{31}$$

B. Unbiased Time-Reversed MMSE-DFE

Now, let us assume that the transmitted data sequence x_n is of a finite length so that the MMSE-DFE can be performed on the time-reversed received signals using the time-reverse of the original channel impulse response [13]. Denoting the time-reversed ISI channel coefficients as $\tilde{h}_n = h^*_{L_h - 1 - n}$, its D-transform is given as $\tilde{h}(D) = D^{L_h - 1} h^*(D^{-*})$. Therefore, the D-transform of the autocorrelation function of the time-reversed channel is given by $R_{\tilde{h}\tilde{h}}(D) = \tilde{h}(D) \tilde{h}^*(D^{-*}) = R_{hh}(D)$. Accordingly, the feedforward and feedback filters of the time-reversed MMSE-DFE, denoted by $\tilde{c}(D)$ and $\tilde{d}(D) - 1$ respectively, are identical to the normal MMSE-DFE filters, i.e.,

$$\tilde{c}(D) = c(D) = \frac{P_x}{P_0 g^*(D^{-*})}, \qquad \tilde{d}(D) = d(D) = g(D). \tag{32}$$

The unbiased output of the time-reversed MMSE-DFE can be expressed similarly to the case of the normal, forward MMSE-DFE except that the unbiased output sequence right after the time-reversed MMSE-DFE should also be time-reversed, in order to get the unbiased equalized output $Y_b(D)$ matched to the input sequence $x(D)$. Therefore,

$$Y_b(D) = x(D) + \frac{P_0}{P_0 - N_0} e_b'(D) \tag{33}$$

where

$$e_b'(D) \triangleq \frac{N_0}{P_0} \left(1 - \frac{1}{g(D)} \right) x(D) + \frac{P_x}{P_0} \left(\frac{w'(D)}{g(D)} \right). \tag{34}$$

Then, the MSE and SNR of the unbiased time-reversed MMSE-DFE are given by

$$\text{MSE}_{URDFE} = \left(\frac{P_0}{P_0 - N_0} \right)^2 \text{E}(|e_{b,n}'|^2) = \frac{P_x N_0}{P_0 - N_0} \tag{35}$$

$$\text{SNR}_{URDFE} \triangleq \frac{P_x}{\text{MSE}_{URDFE}} = \frac{P_0 - N_0}{N_0}. \tag{36}$$

C. Unbiased BiDFE

The structure of the BiDFE is shown in Fig. 1. If we assume that the feedback sequence is correct, the outputs of two unbiased DFEs are:

$$Y_{f,n} = X_n + V_{f,n} \tag{37}$$

$$Y_{b,n} = X_n + V_{b,n} \tag{38}$$

where $V_{f,n}$ and $V_{b,n}$ have D-transforms $V_f(D)$ and $V_b(D)$ as given by (from (28), (29), (33), and (34))

$$V_f(D) = \frac{N_0}{P_0 - N_0} \left(1 - \frac{1}{g^*(D^{-*})} \right) x(D) + \frac{P_x}{P_0 - N_0} \left(\frac{w'(D)}{g^*(D^{-*})} \right) \tag{39}$$

$$V_b(D) = \frac{N_0}{P_0 - N_0} \left(1 - \frac{1}{g(D)} \right) x(D) + \frac{P_x}{P_0 - N_0} \left(\frac{w'(D)}{g(D)} \right). \tag{40}$$

Assuming stationary random processes, we drop time index n for notational simplicity and write: $Y_f = X + V_f$ and $Y_b = X + V_b$. From (30) and (35), the variance of V_f and V_b are also given as:

$$\text{Var}(V_f) = \text{Var}(V_b) = \frac{P_x N_0}{P_0 - N_0}.$$

The variables V_f and V_b are correlated with the correlation coefficient given by

$$
\begin{aligned}
\rho &\triangleq \frac{\mathrm{E}(V_f V_b^*)}{\sqrt{\mathrm{Var}(V_f)\mathrm{Var}(V_b)}} \\
&= \frac{P_0 - N_0}{P_x N_0} \mathrm{E}\left[V_f(D)V_b^*(D^{-*})\right]_0 \\
&= \frac{P_x}{N_0(P_0 - N_0)} \mathrm{E}\left[\left(\frac{1}{g^*(D^{-*})}\right)^2 w'(D)w'^*(D^{-*})\right]_0 \\
&= \frac{P_0^2}{P_x N_0(P_0 - N_0)}\left[\{c(D)\}^2 R_{w'w'}(D)\right]_0 \\
&= \frac{P_0^2}{P_x(P_0 - N_0)}\left[\{c(D)\}^2 R_{hh}(D)\right]_0
\end{aligned}
$$
(41)

(42)

where $[z(D)]_0 = z_0$ with $z(D) = \sum_k z_k D^k$. The equality in (41) holds due to the assumption that X_n is an i.i.d random variable and the self-interference term is removed from the expression $1 - 1/g^*(D^{-*})$.

Since $\mathrm{Var}(V_f) = \mathrm{Var}(V_b)$, the linear MMSE combiner of [7], [14] becomes $Y = \frac{1}{2}(Y_f + Y_b)$. Naturally, the MSE and SNR of the unbiased BiDFE are given as

$$
\mathrm{MSE}_{UBiDFE} = \frac{(1 + \mathrm{Re}[\rho])}{2}\mathrm{MSE}_{UDFE} = \frac{(1 + \mathrm{Re}[\rho])P_x N_0}{2(P_0 - N_0)}
$$
(43)

$$
\mathrm{SNR}_{UBiDFE} \triangleq \frac{P_x}{\mathrm{MSE}_{UBiDFE}} = \frac{2}{(1 + \mathrm{Re}[\rho])}\mathrm{SNR}_{UDFE} = \frac{2(P_0 - N_0)}{(1 + \mathrm{Re}[\rho])N_0}
$$
(44)

where $\mathrm{Re}[\rho]$ denotes the real part of ρ.

Note that the infinite-length normal/time-reversed MMSE-DFE and BiDFE analyzed here do not exploit the *a priori* information of X_n. In other words, the feedforward and feedback filters of DFE are derived by assuming $\mathrm{E}(X_n) = 0$ for all n, meaning that the calculated SNR performance would reflect the non-turbo ideal-decision BiDFE performance with time-invariant filter taps of Section III-C.

V. DERIVATION OF ITERATIVE BiDFE ALGORITHM

We now discuss an iterative BiDFE algorithm. Iterative equalization schemes based on BiDFE are shown in Fig. 2. Basically, the channel equalizer is a SISO equalizer which employs the normal forward DFE, the time-reversed DFE and an LLR combining block. The received data sequence is equalized in both directions by the two DFEs, and the extrinsic information from two DFEs are combined and passed to the error correction code decoder. We show that a proper combining of the two sets of extrinsic information can suppress error propagation and noise further and generate more reliable extrinsic information for the outer decoder.

A. Combining Extrinsic Information

Similarly to the finite-length time-varying feedforward and feedback filter of the normal DFE at time index n, which are previously defined as $\mathbf{c}_n$ in (2) and $\mathbf{d}_n$ in (3), we also define the finite-length time-varying feedforward and feedback filter of the time-reversed DFE at time index n as $\tilde{\mathbf{c}}_n$ and $\tilde{\mathbf{d}}_n$ with the same lengths as $\mathbf{c}_n$ and $\mathbf{d}_n$ respectively. Note that $\tilde{\mathbf{c}}_n$ and $\tilde{\mathbf{d}}_n$ are defined in a similar way as (2) and (3) except that the channel convolution matrix $\tilde{\mathbf{H}}$ for the time-reversed channel is given as

$$\tilde{\mathbf{H}} \triangleq \begin{bmatrix} h_0\, h_1 \cdots h_{L_h-1} & 0 & \cdots & & 0 \\ 0\; h_0\; h_1 & \cdots & h_{L_h-1}\; 0 \cdots & & 0 \\ \ddots & & \ddots & & \ddots \\ 0\; 0\; \cdots & 0 & & h_0 & h_1 \cdots h_{L_h-1} \end{bmatrix}.$$

The unbiased equalizer output [12] corresponding to the transmitted coded symbol from the the normal (forward) and the time-reversed (backward) DFE can be represented respectively as

$$Y_{f,n} = X_n + I_{f,n} + V_{f,n} \tag{45}$$

$$Y_{b,n} = X_n + I_{b,n} + V_{b,n} \tag{46}$$

where $X_n \triangleq x_n$, $V_{f,n} \triangleq v_{f,n}/p_{\{n,0\}}$ and $I_{f,n} \triangleq i_{f,n}/p_{\{n,0\}}$. Also, $V_{b,n} \triangleq v_{b,n}/\tilde{p}_{\{n,0\}}$ and $I_{b,n} \triangleq i_{b,n}/\tilde{p}_{\{n,0\}}$ where $v_{b,n}$ and $i_{b,n}$ are defined similarly to the normal DFE and $\tilde{p}_{\{n,0\}} = \tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}}$ where $\tilde{\mathbf{s}} \triangleq \tilde{\mathbf{H}}[\mathbf{0}_{1 \times L_d}, 1, \mathbf{0}_{1 \times L_c}]^T$. For notational simplicity, we further drop time index n with an understanding that processing remains identical as n progresses: $Y_f = X + I_f + V_f$ and $Y_b = X + I_b + V_b$.

Now, we discuss the problem of how to combine the extrinsic information from two DFEs. Initially, let us consider two unbiased equalizer outputs, which are corrupted by AWGN, corresponding to the transmitted coded symbol X:

$$Y_f = X + U_f$$

$$Y_b = X + U_b$$

where the noise U_f and U_b are assumed to be zero mean Gaussian random variables which are independent of the coded data X but correlated with each other with correlation coefficient ρ.

In order to combine the extrinsic information, it is beneficial to whiten the noise U_f and U_b before combining. The noise correlation matrix $\mathbf{R}$ is defined as

$$\mathbf{R} \triangleq \begin{bmatrix} \mathrm{Var}(U_f) & \mathrm{E}(U_f U_b) \\ \mathrm{E}(U_f U_b) & \mathrm{Var}(U_b) \end{bmatrix} = \begin{bmatrix} N_f & \rho\sqrt{N_f N_b} \\ \rho\sqrt{N_f N_b} & N_b \end{bmatrix}$$

where $N_f \triangleq \mathrm{Var}(U_f)$ and $N_b \triangleq \mathrm{Var}(U_b)$. Then, the eigenvalues of the noise correlation matrix, λ_1 and λ_2, with their corresponding normalized eigenvectors $\mathbf{g}_1$ and $\mathbf{g}_2$ are given by

$$\lambda_1 = \frac{(N_f + N_b) + \sqrt{(N_f - N_b)^2 + 4\rho^2 N_f N_b}}{2}$$

$$\lambda_2 = \frac{(N_f + N_b) - \sqrt{(N_f - N_b)^2 + 4\rho^2 N_f N_b}}{2}$$

$$\mathbf{g}_1 = \frac{1}{\sqrt{g_{11}^2 + g_{21}^2}} \begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix}, \ \mathbf{g}_2 = \frac{1}{\sqrt{g_{12}^2 + g_{22}^2}} \begin{bmatrix} g_{12} \\ g_{22} \end{bmatrix}$$

where $g_{11} = \frac{1}{2}\left[(N_f - N_b) + \sqrt{(N_f - N_b)^2 + 4\rho^2 N_f N_b}\right]$, $g_{12} = \frac{1}{2}\left[(N_f - N_b) - \sqrt{(N_f - N_b)^2 + 4\rho^2 N_f N_b}\right]$, and $g_{21} = g_{22} = \rho\sqrt{N_f N_b}$. It is easy to see that the noise correlation matrix $\mathbf{R}$ is non-singular unless $\rho = \pm 1$. If $\mathbf{R}$ is non-singular, $\mathbf{R}$ can be expanded as $\mathbf{R} = \mathbf{G}\mathbf{\Lambda}\mathbf{G}^{-1}$ where $\mathbf{G} \triangleq [\mathbf{g}_1\ \mathbf{g}_2]$ and $\mathbf{\Lambda} \triangleq \mathrm{Diag}(\lambda_1, \lambda_2)$. Since $\mathbf{G}$ is a unitary matrix, the noise whitening matrix is $\mathbf{A} \triangleq [\mathbf{a}_1\ \mathbf{a}_2] = \mathbf{G}^{-1} = \mathbf{G}^T$ where $\mathbf{a}_1 \triangleq [a_{11}\ a_{21}]^T$ and $\mathbf{a}_2 \triangleq [a_{12}\ a_{22}]^T$. So, given the equalized output vector $\mathbf{Y} \triangleq [Y_f, Y_b]^T$, the whitened vector is $\mathbf{Y}' \triangleq [Y_f', Y_b']^T = \mathbf{A}\mathbf{Y}$ with the new noise correlation matrix $\mathbf{R}' = \mathbf{A}\mathbf{R}\mathbf{A}^T = \mathbf{\Lambda}$. Finally, the extrinsic information of X can be expressed as

$$\begin{aligned}
L_e(X) &= \ln \frac{\Pr(Y_f, Y_b \mid X = +1)}{\Pr(Y_f, Y_b \mid X = -1)} \\
&= \ln \frac{\Pr(Y_f', Y_b' \mid X = +1)}{\Pr(Y_f', Y_b' \mid X = -1)} \\
&= \ln \frac{\Pr(Y_f' \mid X = +1)}{\Pr(Y_f' \mid X = -1)} + \ln \frac{\Pr(Y_b' \mid X = +1)}{\Pr(Y_b' \mid X = -1)} \\
&= \frac{2(a_{11} + a_{12})Y_f'}{\lambda_1} + \frac{2(a_{21} + a_{22})Y_b'}{\lambda_2} \\
&= \frac{2\left(N_b - \rho\sqrt{N_f N_b}\right)Y_f}{(1 - \rho^2)\,N_f N_b} + \frac{2\left(N_f - \rho\sqrt{N_f N_b}\right)Y_b}{(1 - \rho^2)\,N_f N_b} \\
&= \frac{\left(N_b - \rho\sqrt{N_f N_b}\right)}{(1 - \rho^2)\,N_b} L_{e,f}(X) + \frac{\left(N_f - \rho\sqrt{N_f N_b}\right)}{(1 - \rho^2)\,N_f} L_{e,b}(X). \quad (47)
\end{aligned}$$

For the singular noise correlation matrix $\mathbf{R}$ (i.e., $\rho = +1$), $N_f = N_b = N$ and $Y_f = Y_b = Y$ so that $L_{e,f}(X) = L_{e,b}(X)$. Consequently, the extrinsic information of X becomes $L_e(X) = 2Y/N = (L_{e,f}(X) + L_{e,b}(X))/2$. Note that the mean combiner of [9], $L_e(X) = (L_{e,f}(X) + L_{e,b}(X))/2$, can be considered as the proposed combiner with $\rho = +1$. If $\rho = -1$, $U_f = -U_b$ and we can cancel out the noise perfectly by averaging the outputs: $(Y_f + Y_b)/2$. The extrinsic information of X in this case is $L_e(X) = +\infty$ when $(Y_f + Y_b)/2 \geq 0$ while $L_e(X) = -\infty$ when $(Y_f + Y_b)/2 < 0$.

B. Reducing the Combiner Sensitivity to the Estimation Error

Let us consider the effect of errors in estimating ρ on extrinsic information. Write $\hat{\rho} = \rho + \varepsilon$ where ε is the estimation error. Then, the sensitivity of the combiner in (47) to the estimation error can be defined as

$$
\begin{aligned}
\mathrm{S}(\rho) &\triangleq \left| \frac{\partial L_e(X)}{\partial \rho} \right| \\
&= \left| \frac{\left(2\rho N_b - (1 + \rho^2)\sqrt{N_f N_b}\right)}{(1 - \rho^2)^2 \, N_b} L_{e,f}(X) + \frac{\left(2\rho N_f - (1 + \rho^2)\sqrt{N_f N_b}\right)}{(1 - \rho^2)^2 \, N_f} L_{e,b}(X) \right|
\end{aligned}
$$

which approaches infinity as $\rho \to \pm 1$. This means that the combiner of (47) is unfortunately very sensitive to the correlation estimator error, as the magnitude of the correlation becomes large.

The sensitivity of the combiner can be reduced if we assume that the variance of U_f and U_b are the same, i.e., $N = N_f = N_b = (N_f + N_b)/2$. This assumption is reasonable when the same feedforward and feedback filter length is used in both DFEs. Then, from (47), the combined extrinsic information of X for non-singular $\mathbf{R}$ is simply given as

$$
L_e(X) = \frac{1}{(1 + \rho)} \Big(L_{e,f}(X) + L_{e,b}(X) \Big) \tag{48}
$$

with the sensitivity to the correlation estimation error

$$
\mathrm{S}(\rho) = \left| \frac{1}{(1 + \rho)^2} \Big(L_{e,f}(X) + L_{e,b}(X) \Big) \right|.
$$

Although the sensitivity of this combiner to the estimation error also goes to infinity as $\rho \to -1$, it shows more robustness as $\rho \to +1$ since $\lim_{\rho \to +1} \mathrm{S}(\rho) = |(L_{e,f}(X) + L_{e,b}(X))/4|$.

C. Application to the BiDFE Algorithm

In this paper, although the composite noise $I_{f,n} + V_{f,n}$ and $I_{b,n} + V_{b,n}$ are not Gaussian, we exploit the combiner of (48) in order to produce the combined extrinsic information to be passed to the convolutional decoder. The noise correlation coefficient between $I_{f,n} + V_{f,n}$ and $I_{b,n} + V_{b,n}$ is naturally defined as

$$
\rho_n \triangleq \frac{\mathrm{E}\left\{ (I_{f,n} - \mathrm{E}(I_{f,n}) + V_{f,n})\,(I_{b,n} - \mathrm{E}(I_{b,n}) + V_{b,n}) \right\}}{\sqrt{(\mathrm{Var}(I_{f,n}) + \mathrm{Var}(V_{f,n}))\,(\mathrm{Var}(I_{b,n}) + \mathrm{Var}(V_{b,n}))}}. \tag{49}
$$

Unfortunately, it is difficult to compute the correlation coefficient analytically in the presence of decision feedback errors. However, assuming that the noise is stationary, we have $\rho_n = \rho$ and the correlation coefficient can be estimated through time-averaging:

$$
\hat{\rho} = \frac{\sum \left\{ (Y_{f,n} - \hat{X}_{f,n} - \mathrm{E}(I_{f,n}))(Y_{b,n} - \hat{X}_{b,n} - \mathrm{E}(I_{b,n})) \right\}}{\sqrt{\sum (Y_{f,n} - \hat{X}_{f,n} - \mathrm{E}(I_{f,n}))^2}\sqrt{\sum (Y_{b,n} - \hat{X}_{b,n} - \mathrm{E}(I_{b,n}))^2}} \tag{50}
$$

where the summations are over some reasonably large finite window. Note that the hard decisions for the transmitted symbols in normal and time-reversed DFEs might be different; in estimating the correlation coefficient, we only consider those noise samples for which $\hat{X}_{f,n}$ and $\hat{X}_{b,n}$ are identical.

Let us summarize our LLR combining method: 1) The extrinsic information $L_{e,f}(X_n)$ and $L_{e,b}(X_n)$ for $n = 1, 2, \ldots, L$ are acquired according to (20) in the normal and time-reversed MMSE-DFE settings. 2) Estimate the noise correlation coefficient, $\hat{\rho}$, between $I_{f,n} + V_{f,n}$ and $I_{b,n} + V_{b,n}$ by (50). 3) Generate the combined extrinsic information $L_e(X_n)$ according to (48) with $\rho_n = \hat{\rho}$.

D. Correlation Analysis under Ideal Feedback

We provide correlation analysis in the following. The analysis will allow validation of (50) in different scenarios. The observation of how the simulated correlation coefficient (50) converges to the analytically computed one under the assumptions of ideal feedback and perfect *a priori* information will also provide useful insights into the iterative behaviour of the proposed turbo BiDFE.

First of all, the noise variance of $V_{f,n}$ and $V_{b,n}$ from the time-varying filters are:

$$\mathrm{Var}(V_{f,n}) = (1 - \mathbf{s}^T \mathbf{c}_n)/\mathbf{c}_n^T \mathbf{s}$$

$$\mathrm{Var}(V_{b,n}) = (1 - \tilde{\mathbf{s}}^T \tilde{\mathbf{c}}_n)/\tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}}.$$

When we assume ideal decision feedback, $\mathrm{Pr}(I_f = 0) = \mathrm{Pr}(I_b = 0) = 1$ so that $I_{f,n} = I_{b,n} = 0$, the noise correlation coefficient ρ_n between $V_{f,n}$ and $V_{b,n}$ becomes

$$
\begin{aligned}
\rho_n &\triangleq \frac{\mathrm{E}(V_{f,n}V_{b,n})}{\sqrt{\mathrm{Var}(V_{f,n})\mathrm{Var}(V_{b,n})}} \\
&= \frac{\mathrm{E}\left[\left\{\dfrac{1}{p_{\{n,0\}}} \sum\limits_{j=0}^{L_c} c_{\{n,j\}} w_{n+j}\right\}\left\{\dfrac{1}{\tilde{p}_{\{n,0\}}} \sum\limits_{k=0}^{L_c} \tilde{c}_{\{n,k\}} w_{n-k+L_h-1}\right\}\right]}{\sqrt{(1 - \mathbf{s}^T \mathbf{c}_n)/\mathbf{c}_n^T \mathbf{s}}\sqrt{(1 - \tilde{\mathbf{s}}^T \tilde{\mathbf{c}}_n)/\tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}}}} \quad (51) \\
&= \frac{\sum\limits_{j=0}^{L_c} \sum\limits_{k=0}^{L_c} c_{\{n,j\}} \tilde{c}_{\{n,k\}} \mathrm{E}\left[w_{n+j} w_{n-k+L_h-1}\right]}{\sqrt{\mathbf{c}_n^T \mathbf{s}(1 - \mathbf{s}^T \mathbf{c}_n)}\sqrt{\tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}}(1 - \tilde{\mathbf{s}}^T \tilde{\mathbf{c}}_n)}} \\
&= N_0 \left(\frac{\sum\limits_{j=0}^{L_c} \sum\limits_{k=0}^{L_c} c_{\{n,j\}} \tilde{c}_{\{n,k\}} \delta(j + k + 1 - L_h)}{\sqrt{\mathbf{c}_n^T \mathbf{s}(1 - \mathbf{s}^T \mathbf{c}_n)}\sqrt{\tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}}(1 - \tilde{\mathbf{s}}^T \tilde{\mathbf{c}}_n)}}\right) \quad (52)
\end{aligned}
$$

where $\delta(t)$ is defined as: if $t = 0$, $\delta(t) = 1$; otherwise, $\delta(t) = 0$. The equality in (51) holds because X_n is an i.i.d random variable.

If the time-invariant filters are used instead of the time-varying filters, the variances of $V_{f,n}$ and $V_{b,n}$ become

$$\text{Var}(V_{f,n}) = \mathbf{c}^T \left(\mathbf{H}\boldsymbol{\Sigma}_n\mathbf{H}^T - z_n\mathbf{s}\mathbf{s}^T + N_0\mathbf{I} \right) \mathbf{c}/ \left(\mathbf{c}^T\mathbf{s} \right)^2$$

$$\text{Var}(V_{b,n}) = \tilde{\mathbf{c}}^T \left(\tilde{\mathbf{H}}\tilde{\boldsymbol{\Sigma}}_n\tilde{\mathbf{H}}^T - z_n\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T + N_0\mathbf{I} \right) \tilde{\mathbf{c}}/ \left(\tilde{\mathbf{c}}^T\tilde{\mathbf{s}} \right)^2 .$$

Then, the noise correlation coefficient can be also obtained as

$$\rho_n = N_0 \left(\frac{\sum\limits_{j=0}^{L_c}\sum\limits_{k=0}^{L_c} c_j\tilde{c}_k\delta(j+k+1-L_h)}{\sqrt{\mathbf{c}^T \left(\mathbf{H}\boldsymbol{\Sigma}_n\mathbf{H}^T - z_n\mathbf{s}\mathbf{s}^T + N_0\mathbf{I} \right) \mathbf{c}}\sqrt{\tilde{\mathbf{c}}^T(\tilde{\mathbf{H}}\tilde{\boldsymbol{\Sigma}}_n\tilde{\mathbf{H}}^T - z_n\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T + N_0\mathbf{I})\tilde{\mathbf{c}}}} \right) . \tag{53}$$

Now, let us consider some special cases.

1) No A Priori Information: When no *a priori* information is available, i.e., $\text{E}(X_n) = 0$ for all n, the feedforward and feedback filters are the same as the time-invariant filters and the noise variances are stationary:

$$\text{Var}(V_{f,n}) = \text{Var}(V_f) = (1 - \mathbf{s}^T\mathbf{c})/\mathbf{c}^T\mathbf{s}$$

$$\text{Var}(V_{b,n}) = \text{Var}(V_b) = (1 - \tilde{\mathbf{s}}^T\tilde{\mathbf{c}})/\tilde{\mathbf{c}}^T\tilde{\mathbf{s}}.$$

Therefore, the noise correlation coefficient is given by

$$\rho_n = \rho = N_0 \left(\frac{\sum\limits_{j=0}^{L_c}\sum\limits_{k=0}^{L_c} c_j\tilde{c}_k\delta(j+k+1-L_h)}{\sqrt{\mathbf{c}^T\mathbf{s}(1 - \mathbf{s}^T\mathbf{c})}\sqrt{\tilde{\mathbf{c}}^T\tilde{\mathbf{s}}(1 - \tilde{\mathbf{s}}^T\tilde{\mathbf{c}})}} \right) . \tag{54}$$

We observed that the noise correlation coefficient of the infinite-length BiDFE in (42) is almost identical to that of the finite-length BiDFE in (54) when L_c is chosen to be long enough.

2) Time-varying Filters with Perfect A Priori Information: When several iterations are performed at high SNRs in turbo equalization, the perfect *a priori* information could be available, i.e., $\text{E}(X_n) = X_n$ for all n. When $\text{E}(X_n) = X_n$ for all n, the feedforward filters $\mathbf{c}_n$ and $\tilde{\mathbf{c}}_n$ of two DFEs become the normalized matched filters corresponding to the forward and reverse channel impulse responses:

$$\mathbf{c}_n = A\left[h_0, h_1, \ldots, h_{L_h-1}, \mathbf{0}_{1\times L_c-L_h+1}\right]^T$$

$$\tilde{\mathbf{c}}_n = A\left[h_{L_h-1}, h_{L_h-2}, \ldots, h_0, \mathbf{0}_{1\times L_c-L_h+1}\right]^T$$

where A is a real-valued constant depending on SNR, i.e., $A \triangleq 1/(N_0 + \sum_{k=0}^{L_h-1} |h_k|^2)$. Moreover, since the first terms of $V_{f,n}$ and $V_{b,n}$ disappear, the noise variances are simply:

$$\mathrm{Var}(V_{f,n}) = \mathrm{Var}(V_f) = N_0 \mathbf{c}_n^T \mathbf{c}_n / (\mathbf{c}_n^T \mathbf{s})^2 = \frac{N_0 A^2}{(\mathbf{c}_n^T \mathbf{s})^2} \sum_{k=0}^{L_h-1} |h_k|^2$$

$$\mathrm{Var}(V_{b,n}) = \mathrm{Var}(V_b) = N_0 \tilde{\mathbf{c}}_n^T \tilde{\mathbf{c}}_n / (\tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}})^2 = \frac{N_0 A^2}{(\tilde{\mathbf{c}}_n^T \tilde{\mathbf{s}})^2} \sum_{k=0}^{L_h-1} |h_k|^2.$$

Accordingly, the noise correlation coefficient is

$$\rho_n = \rho = 1. \tag{55}$$

Note that the noise correlation coefficient ρ with perfect *a priori* information converges to 1 regardless of the SNR value. As will be shown shortly, the measured correlation coefficient using simulated turbo BiDFE outputs indeed approaches 1, as turbo iteration progresses. This indicates that both assumptions - ideal decision feedback and perfect *a priori* information - are reasonable.

3) Time-invariant Filters with Perfect A Priori Information: When the time-invariant filters are used with perfect *a priori* information, the time-invariant DFEs yield the noise variances as

$$\mathrm{Var}(V_{f,n}) = \mathrm{Var}(V_f) = N_0 \mathbf{c}^T \mathbf{c} / (\mathbf{c}^T \mathbf{s})^2$$

$$\mathrm{Var}(V_{b,n}) = \mathrm{Var}(V_b) = N_0 \tilde{\mathbf{c}}^T \tilde{\mathbf{c}} / (\tilde{\mathbf{c}}^T \tilde{\mathbf{s}})^2.$$

The noise correlation coefficient is also simply given by

$$\rho_n = \rho = \frac{\sum_{j=0}^{L_c} \sum_{k=0}^{L_c} c_j \tilde{c}_k \delta(j + k + 1 - L_h)}{\sqrt{\mathbf{c}^T \mathbf{c}} \sqrt{\tilde{\mathbf{c}}^T \tilde{\mathbf{c}}}}. \tag{56}$$

As will be discussed in the next section, in the simulation of turbo BiDFE with time-invariant taps it is observed that the BiDFE output correlation does indeed converge to (56), indicating again that the assumptions of error-free decisions and perfect *a priori* information are reasonable.

VI. SIMULATION RESULTS

In this section, simulation results of several iterative equalization schemes are presented. The transmitted symbols are encoded with a recursive rate-$1/2$ convolutional code encoder with parity generator $(1 + D^2)/(1 + D + D^2)$ with 2^{11} message bits and are modulated by binary phase-shift keying (BPSK) so that $x_n \in \{\pm 1\}$. We also assume that the noise is AWGN, and the noise variance and the channel information are perfectly known to the receiver. The ISI channels with impulse responses $\mathbf{h_1} = (1/\sqrt{19})[1 \quad 2 \quad 3 \quad 2 \quad 1]^T$ and $\mathbf{h_2} = (1/\sqrt{44})[1 \quad 2 \quad 3 \quad 4 \quad 3 \quad 2 \quad 1]^T$ investigated in [3] and

[10] are used for evaluating the performance of the iterative equalizers. These channels are considered very severe ISI channels as the channel spectra possess nulls over the Nyquist band, as shown in Fig. 3. Finally, the decoder is implemented using the BCJR algorithm. Only the SISO equalizer changes from one scheme to another. The MMSE-DFE with 17 feedforward taps and 4 feedback taps is used for both the normal and the time-reversed DFEs on $\mathbf{h_1}$ while MMSE-DFE with 21 feedforward taps and 6 feedback taps is used on $\mathbf{h_2}$. Finally, the linear MMSE equalizer uses 21 taps for $\mathbf{h_1}$ and 27 taps for $\mathbf{h_2}$.

Six different equalizer types are simulated in this work. The notation "TV-" denotes equalizers with time-varying filters while "TIV-" indicates those with time-invariant filters. For instance, "TV-LE" in the legend indicates the linear MMSE equalizer with a time-varying filter. The "Proposed DFE" uses the proposed LLR mapping of (20) while "DFE" uses the conventional LLR mapping (as used in [3]) The "Proposed BiDFE" is the iterative BiDFE algorithm which is described in Section V. In other words, "Proposed BiDFE" uses the the proposed LLR generation for both normal and time-reversed DFEs along with the proposed extrinsic information combiner of (48) in conjunction with the noise correlation coefficient of (50). The "BiDFE (mean combiner)" is the iterative BiDFE algorithm with the conventional LLR mapping and the mean combiner, $L_e(X) = (L_{e,f}(X) + L_{e,b}(X))/2$ (of [9]), simulated for performance comparison purposes. Finally, "MAP" is the optimal equalizer implemented via the BCJR algorithm.

A thorough comparison is given in [3] on the required complexity levels of the SISO-LE, SISO-DFE and the MAP equalizers. The exact level of implementation complexity is hard to assess as it depends highly on specific VLSI architecture details. Roughly speaking, however, it is safe to say that the number of multiplications and additions increases as an exponential function of the channel memory length for the MAP equalizer whereas the number of the same operations is a quadratic function of both the channel memory length and the filter length for the TV-LE and the TV-DFE, as shown in [3]. The number of operations, on the other hand, increases only linearly for the TIV-LE and the TIV-DFE [3]. The BiDFE equalizers, including the proposed BiDFE methods, require roughly twice as many operations as the DFE counterparts, due to the presence of the time-reversed filter components. Most notably, while the complexity of the proposed BiDFE with time-invariant filters is considerably lower than that of the MAP equalizer as well as the TV-LE, the performance is significantly better than the TV-LE.

Fig. 4 shows the performance of several turbo equalizers with time-varying filters after 20 iterations. TV-DFE with the conventional LLR mapping shows poor performance but once the proposed LLR generations are used ("Proposed TV-DFE"), the DFE performance becomes clearly better than the TV-LE method of [3], except at very high SNRs where all schemes other than the conventional DFE perform

comparably. The "Proposed TV-BiDFE" is considerably better than the TV-BiDFE based on the mean combiner, approaching the performance of the MAP scheme.

Fig. 5 shows the BER performance of time-invariant-filter-based turbo equalizers. As the figure indicates, the "Proposed TIV-DFE" also shows superior performance to the "TIV-DFE". The performance of "Proposed TIV-BiDFE" is very close to the performance of the MAP equalizer while requiring low computational complexity based on the use of time-invariant filters. Also notice that both "Proposed TIV-DFE" and "Proposed TIV-BiDFE" achieve decision-error-free performance at low BERs, indicating the error propagation effect has been nearly eliminated using the proposed LLR generation method. It is noteworthy that the proposed BiDFE algorithm still provides near-optimal performance even with the time-invariant filter taps. While the TIV-BiDFE based on the existing mean combiner appears to perform almost as well, the EXIT chart analysis to be discussed below indicate that with a smaller number of turbo iterations, its performance is distinctly inferior to the proposed TIV-BiDFE based on the new combining method.

Figs. 6 and 7 show a similar set of simulation results now applied to the more severe ISI channel h_2. While all DFE-based schemes lag clearly behind the BCJR-based scheme at the error rates simulated, the proposed BiDFE scheme in both the time-varying and time-invariant filter cases outperform the LE scheme by a significant margin. In fact, in this severe channel the BER curve of the LE scheme, even with time-varying filters, appears to diverge considerably from the ideal no-ISI curve. Overall, the proposed BiDFE based on time-invariant filter taps offer excellent performance-complexity trade-off.

The noise correlation in one block of coded data bits is described in Fig. 8, at different iteration numbers at a 6 dB SNR on h_1. The correlation coefficient of "Proposed TV-BiDFE" goes to 1 as the number of iterations increases because the *a priori* information from the decoder becomes reliable, and the time-varying filters in the normal and the time-reversed DFEs produce essentially the same equalized output sequences. This phenomenon of Fig. 8 validates (55). On the other hand, the correlation coefficient of "Proposed TIV-BiDFE" actually decreases as the number of iterations increases, and the noise correlation coefficient converges to that of "TIV-BiDFE with Ideal Feedback" or the correlation coefficient of (56). This is because the decision feedback errors disappear and the perfect *a priori* information is available from decoder. Note that the filter coefficients in both DFEs do not change with the *a priori* information.

In general, it is quite difficult to analyse the iterative equalization and decoding schemes. We rely on the oft-used extrinsic information transfer (EXIT) chart of [15] to develop insights into the convergence behaviour of the turbo equalizers. The EXIT chart is a diagram demonstrating the mutual information (MI) transfer characteristics of the two constituent modules which exchange soft information. In the EXIT

charts, the behavior of the channel equalizer is described with its input and output on the horizontal and vertical axis, respectively, while the behavior of the decoder is described in opposite way. The pair of EXIT chart curves typically defines a path for the MI trajectory to move up during iterative processing of soft information. The number of stairs that a given MI trajectory takes to reach the highest value indicates the necessary number of iterations toward convergence.

Figs. 9 and 11 show the EXIT chart corresponding to time-varying-filter-based equalizers for $\mathbf{h_1}$ at a 6 dB SNR and $\mathbf{h_2}$ at a 10 dB SNR while Figs. 10 and 12 show the similar EXIT charts for time-invariant-filter-based schemes. Although not shown here to avoid excessive cluttering, the trajectories of "TV-DFE" and "TIV-DFE" move up for the first couple of iterations, but then quickly fizzle out due to the inadequate extrinsic LLR generations that cannot handle error propagation. However, the trajectories of "Proposed TV-DFE" and "Proposed TIV-DFE" keep moving up as the number of iterations increases, clearly indicating the advantage and effectiveness of the proposed LLR generation method. However, the trajectory of "Proposed TIV-DFE" at 6 dB or 10 dB does not reach the maximum possible value since the filters do not fully exploit the *a priori* information from the decoder. The trajectories of the "Proposed TV-BiDFE" and "Proposed TIV-BiDFE" indicate that these schemes move from 0 bit of mutual information to 1 bit with a less number of iteration runs than "Proposed TV-DFE", "Proposed TIV-DFE", "TV-LE", or "TIV-LE".

We notice, however, that the proposed BiDFE scheme requires more iterations in achieving the full performance, relative to the MAP equalizer (whose trajectory is not shown to avoid cluttering). Nevertheless, the proposed BiDFE method offers a reasonable tradeoff among complexity, performance, and latency.

Finally, Fig. 13 shows the SNR comparison at the output of the unbiased DFE and BiDFE assuming ideal feedback on the channel $\mathbf{h_1}$ when the *a priori* information is not available. As the figure shows, the output SNR of BiDFE is considerably higher than the output SNR of DFE but with a certain gap to the matched filter bound (MFB).

VII. CONCLUSION

In this paper, we proposed new SISO DFE and BiDFE structures well-suited to turbo equalization. The proposed LLR generation designed to reduce error propagation indeed provides decision-error-free performance in the DFE in turbo equalizer setting. When further employing an LLR combining method that estimates the correlation between the forward and backward DFE outputs and whitens them, the resulting performance is remarkably good given the simple structure of the BiDFE, relative to that of

the BCJR equalizer. The proposed LLR generation and combining methods remain effective even when a time-invariance constraint is imposed on the feedforward and feedback filters of the DFEs. Overall, the proposed BiDFE method based on time-invariant filter taps provides the excellent performance-complexity tradeoff for severe ISI channels where the linear SISO equalizer fails to operate adequately.

REFERENCES

[1] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *European Trans. Telecommun.*, vol. 6, no. 5, pp. 507-511, Sep.-Oct. 1995.

[2] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. Information Theory*, vol. IT-20, pp. 284-287, Mar. 1974.

[3] M. Tüchler, R. Kötter, and A. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 754-767, May 2002.

[4] A. Fertner, "Improvement of bit-error-rate in decision feedback equalizer by preventing decision-error propagation," *IEEE Trans. Signal Processing*, vol. 46, no. 7, pp. 1872-1877, Jul. 1998.

[5] S. Ariyavisitakul and Y. Li, "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE Trans. Selected Areas in Communications*, vol. 16, no. 9, pp. 1670-1678, Dec. 1998.

[6] R. R. Lopes and J. R. Barry, "The Soft-Feedback Equalization for Turbo Equalization of Highly Dispersive Channels," *IEEE Trans. Communications*, vol. 54, no. 5, pp. 783-788, May 2006.

[7] J. Balakrishnan and C. R. Johnson, Jr., "Bidirectional Decision Feedback Equalizer: Infinite Length Results," *In Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2001, pp. 1450-1454.

[8] J. Nelson, A. Singer, U. Madhow, and C. McGahey, "BAD: bidirectional arbitrated decision-feedback equalization," *IEEE Trans. Communications*, vol. 53, no. 2, pp. 214-218, Feb. 2005.

[9] J. Jiang, C. He, E. M. Kurtas, and K. R. Narayanan, "Performance of soft feedback equalization over magnetic recording channels," *In Proc. Intermag*, San Diego, CA, May 2006, pp. 795.

[10] J. Moon and F. R. Rad, "Turbo Equalization via Constrained-Delay APP Estimation With Decision Feedback," *IEEE Trans. Communications*, vol. 53, no. 12, pp. 2102-2113, Dec. 2005.

[11] F. R. Rad and J. Moon, "Turbo Equalization Utilizing Soft Decision Feedback," *IEEE Trans. Magnetics*, vol. 41, no. 10, pp. 2998-3000, Oct. 2005.

[12] J. Cioffi, G. Dudevior, M. Eyuboglu, and G. Forney, "MMSE decision-feedback equalizers and coding - Part I and II," *IEEE Trans. Communications*, vol. 43, no. 10, pp. 2582-2604, Oct. 1995.

[13] S. Ariyavisitakul, "A Decision Feedback Equalizer with Time-Reversal Structure," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 599-613, Apr. 1992.

[14] J. Balakrishnan and C. R. Johnson, Jr., "Time-Reversal Diversity in Decision Feedback Equalization," *In Proc. of Allerton Conf. on Comm. Control and Computing*, Monticello, IL, Oct. 2000.

[15] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Trans. Communications*, vol. 49, no. 10, pp. 1727-1737, Oct. 2001.
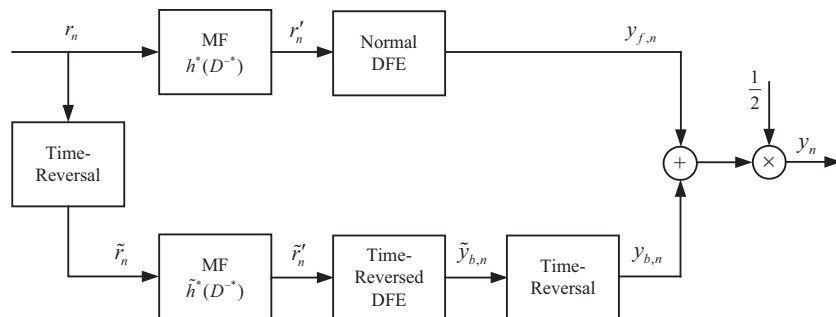
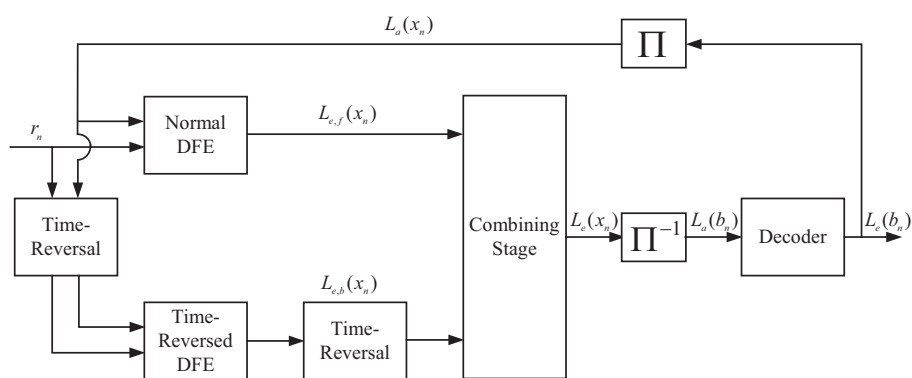Fig. 1: Bidirectional Decision Feedback Equalizer: Infinite Length.

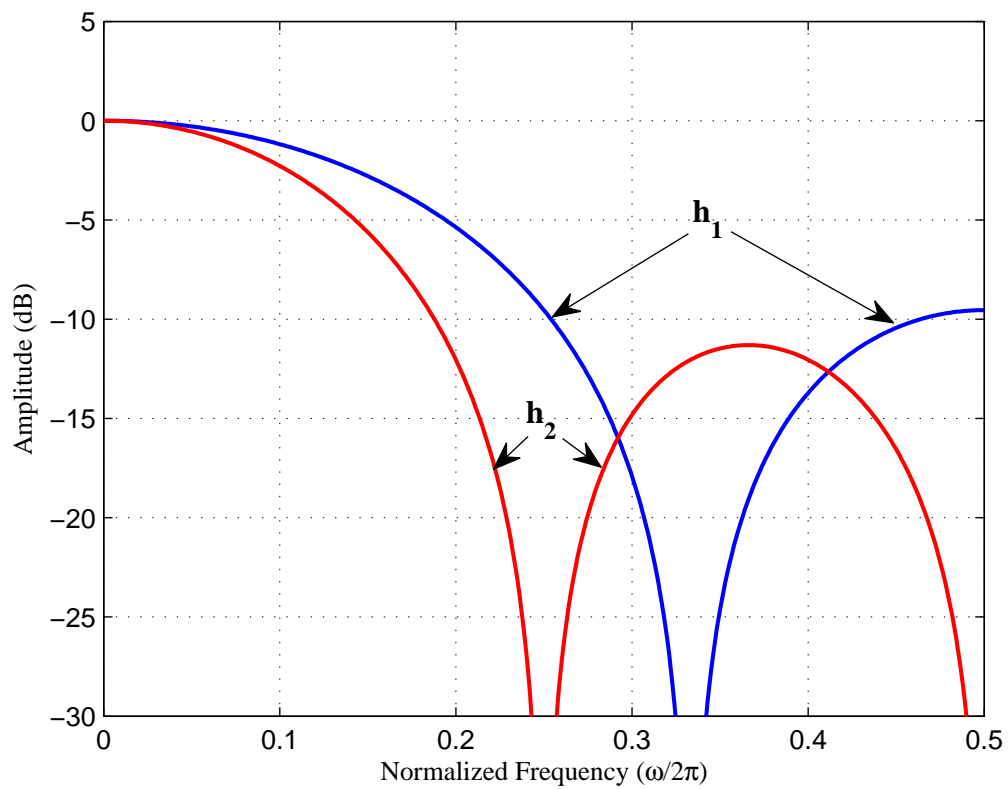Fig. 2: Iterative Equalization Scheme based on BiDFE.

Fig. 3: Frequency Magnitude Response of the ISI Channels: $\mathbf{h_1} = (1/\sqrt{19})[1 \quad 2 \quad 3 \quad 2 \quad 1]^T$, $\mathbf{h_2} = (1/\sqrt{44})[1 \quad 2 \quad 3 \quad 4 \quad 3 \quad 2 \quad 1]^T$.
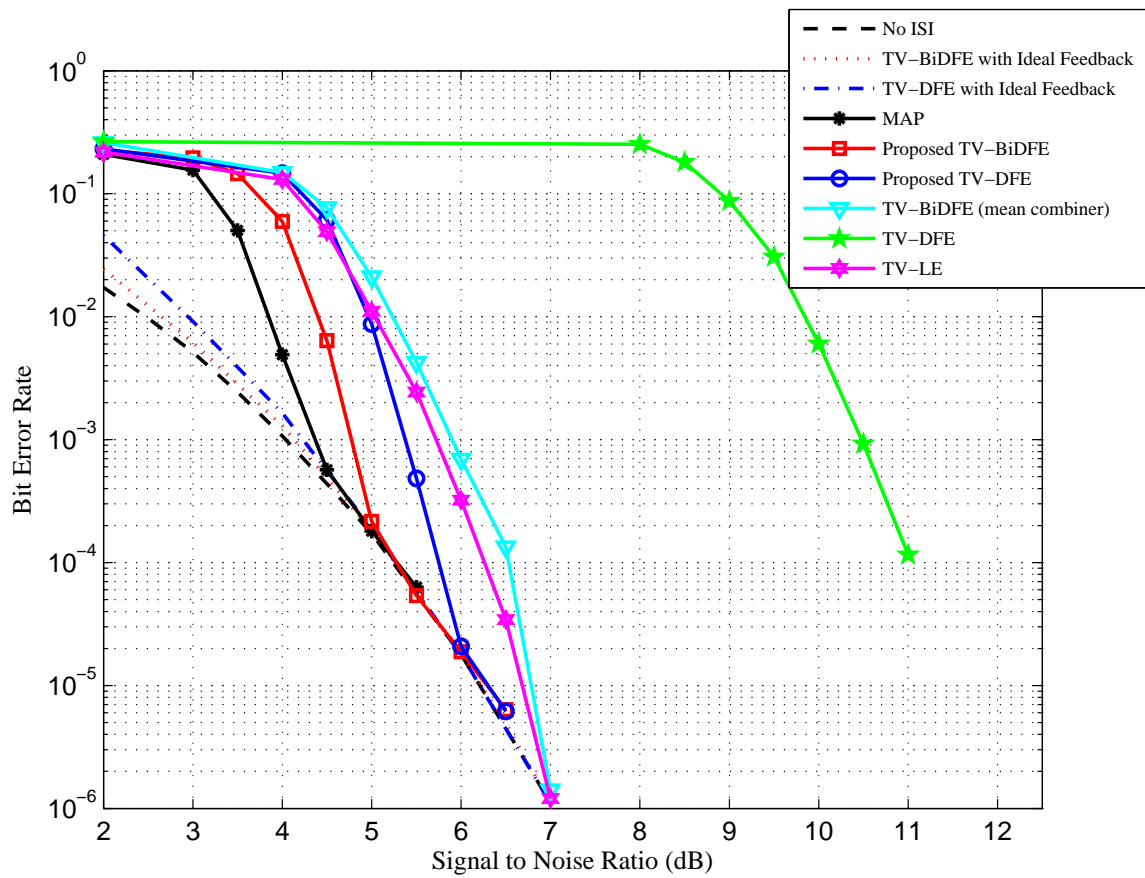
Fig. 4: BER Curve on the Channel $\mathbf{h_1}$ after 20 Iterations with Time-varying Filters.
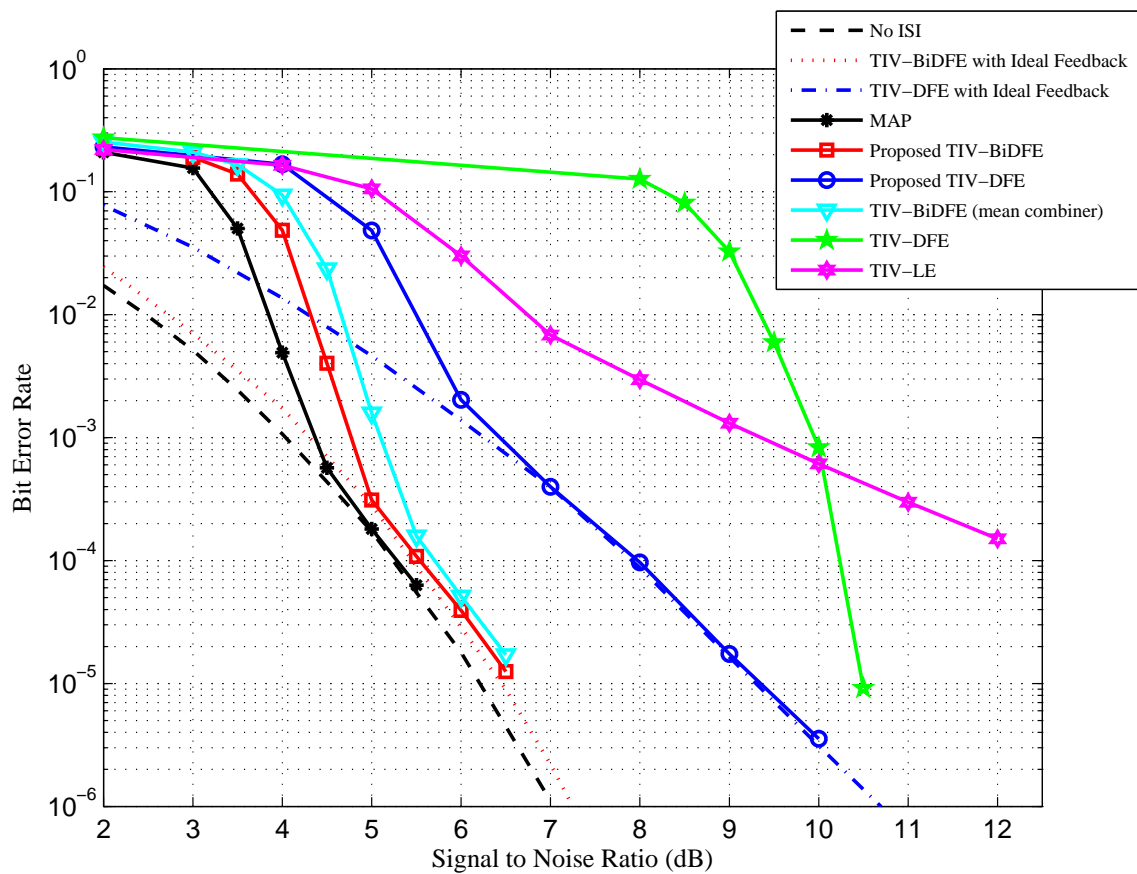
Fig. 5: BER Curve on the Channel $\mathbf{h_1}$ after 20 iterations with Time-invariant Filters.
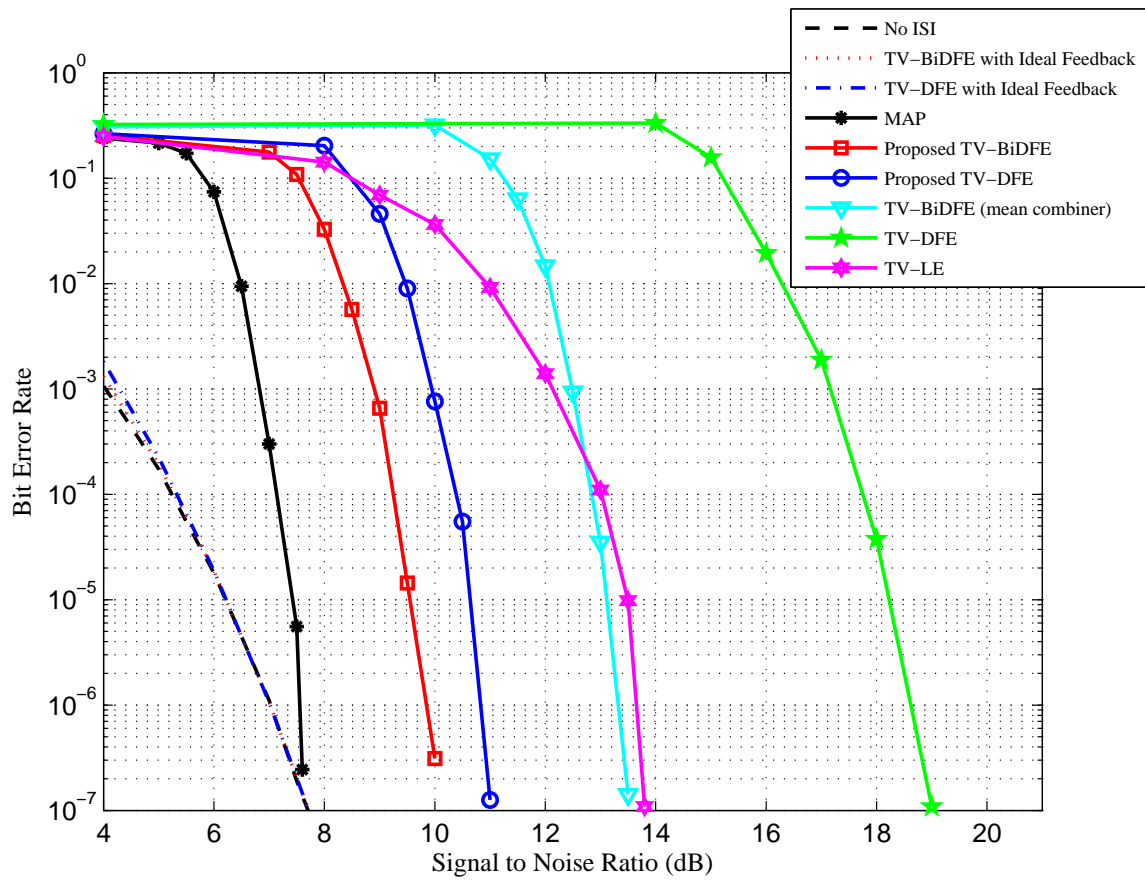
Fig. 6: BER Curve on the Channel $\mathbf{h_2}$ after 20 Iterations with Time-varying Filters.
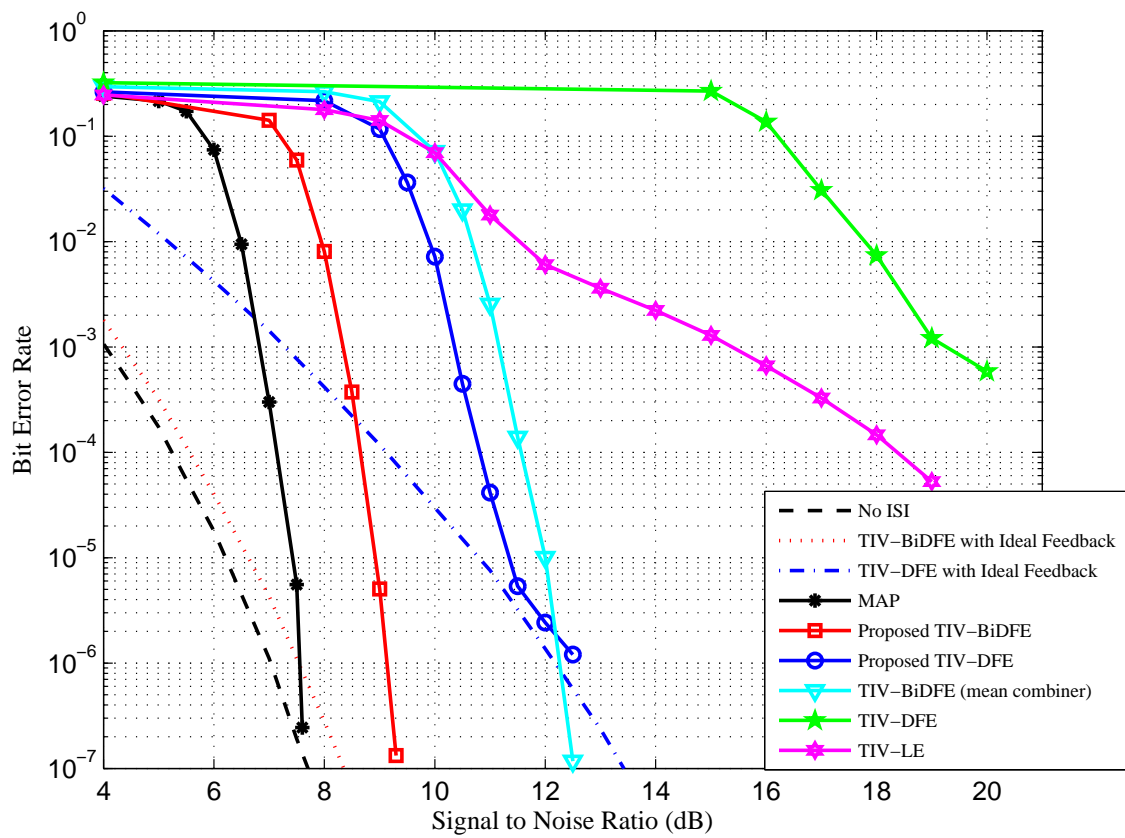
Fig. 7: BER Curve on the Channel $\mathbf{h_2}$ after 20 iterations with Time-invariant Filters.
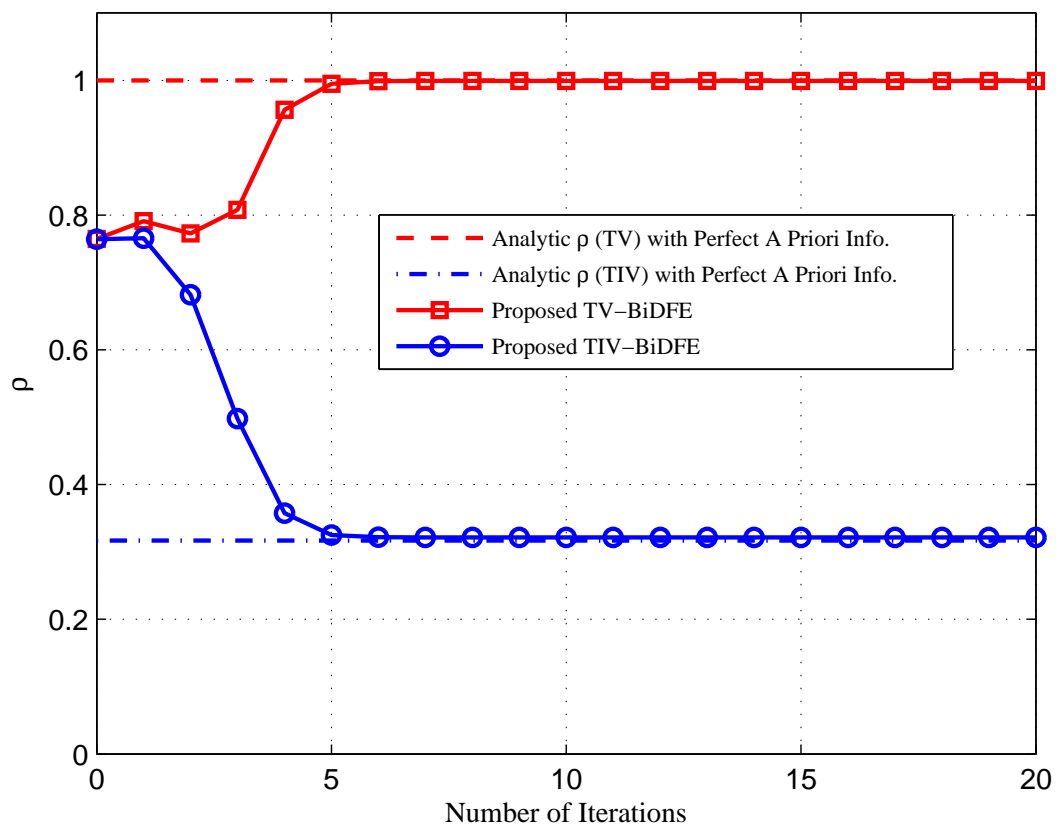
Fig. 8: Noise Correlation of "Proposed BiDFE" on the Channel $\mathbf{h_1}$.
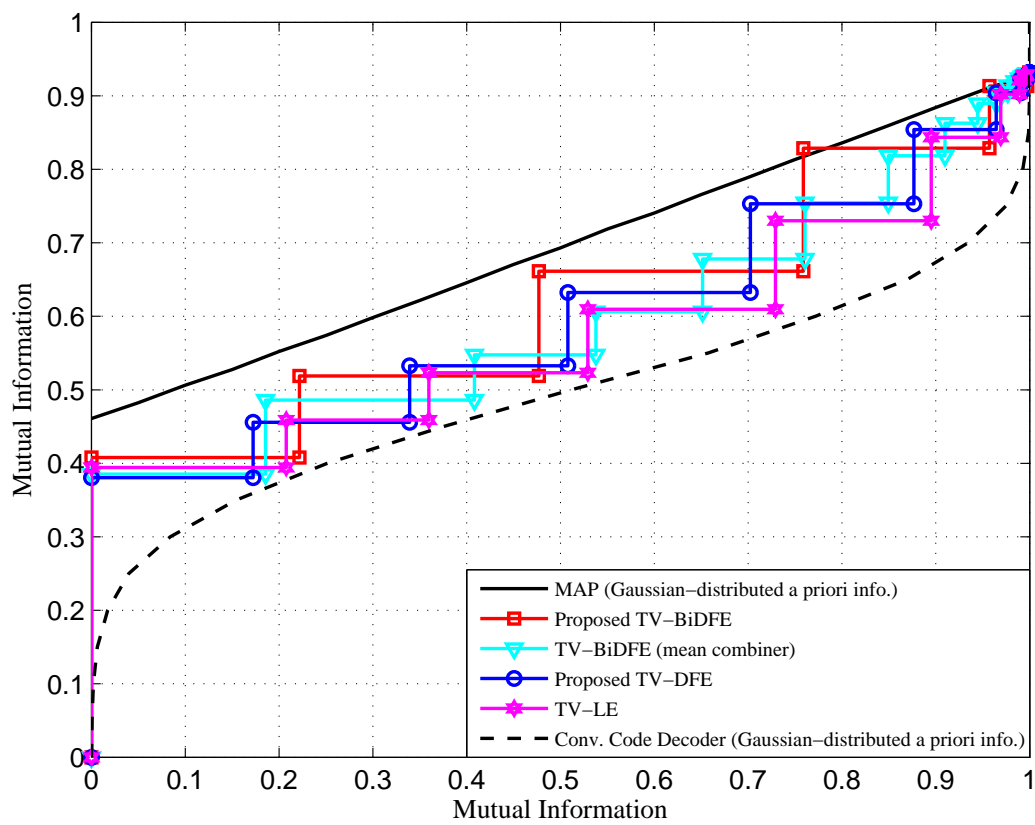
Fig. 9: EXIT Chart on the Channel $\mathbf{h_1}$ at a 6 dB with Time-varying Filters.
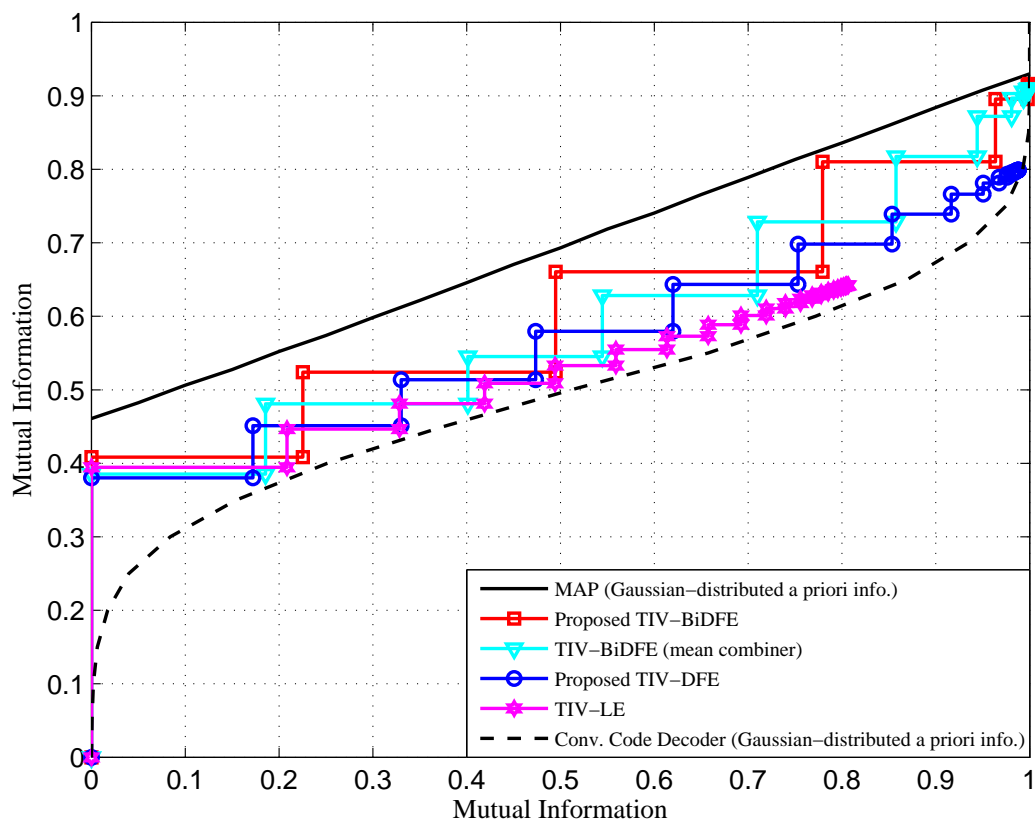
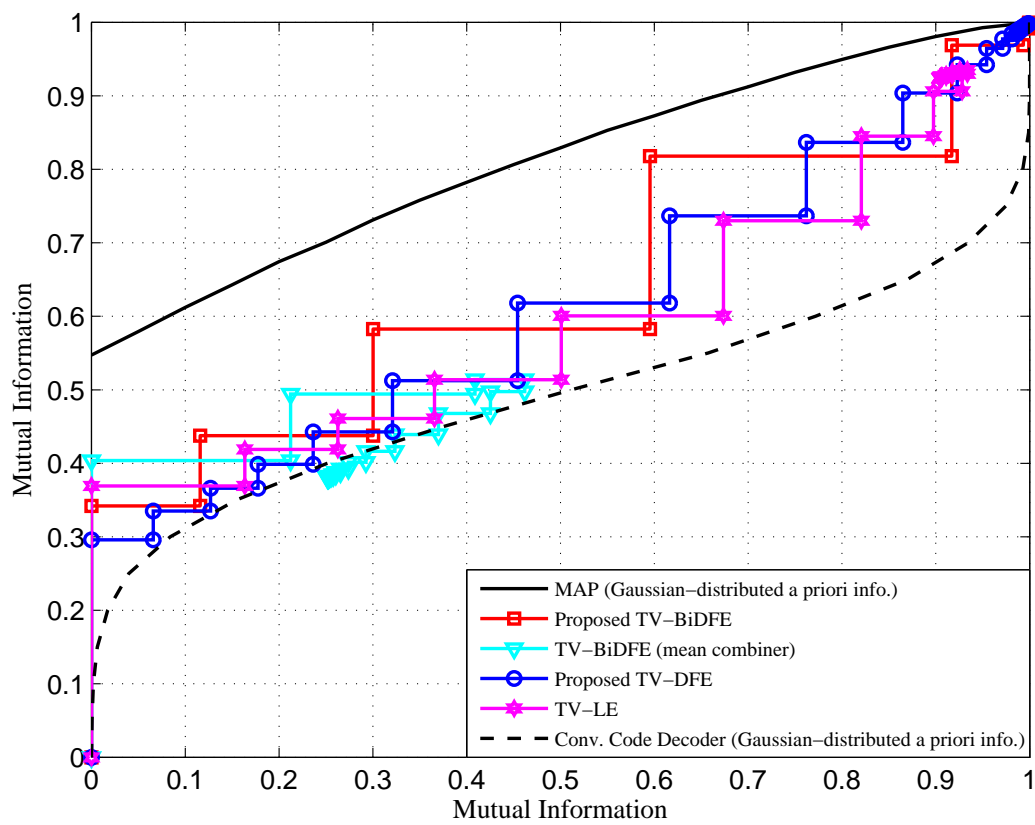Fig. 10: EXIT Chart on the Channel $\mathbf{h_1}$ at a 6 dB with Time-invariant Filters.

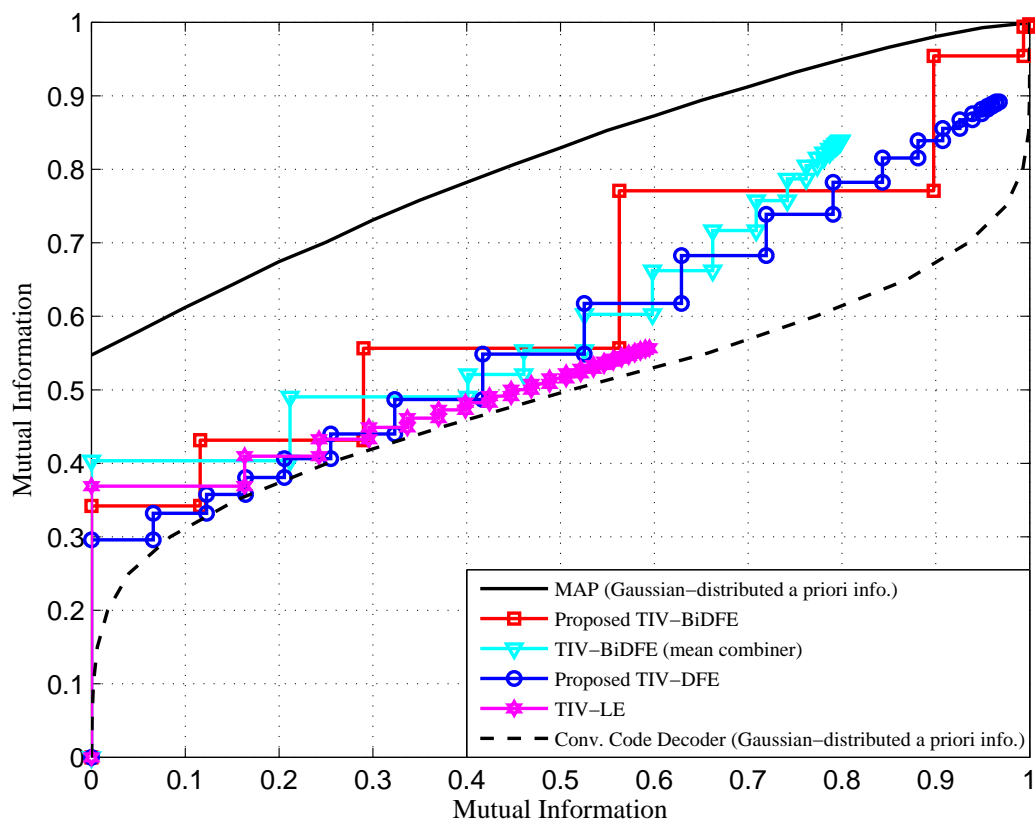Fig. 11: EXIT Chart on the Channel $\mathbf{h_2}$ at a 10 dB with Time-varying Filters.

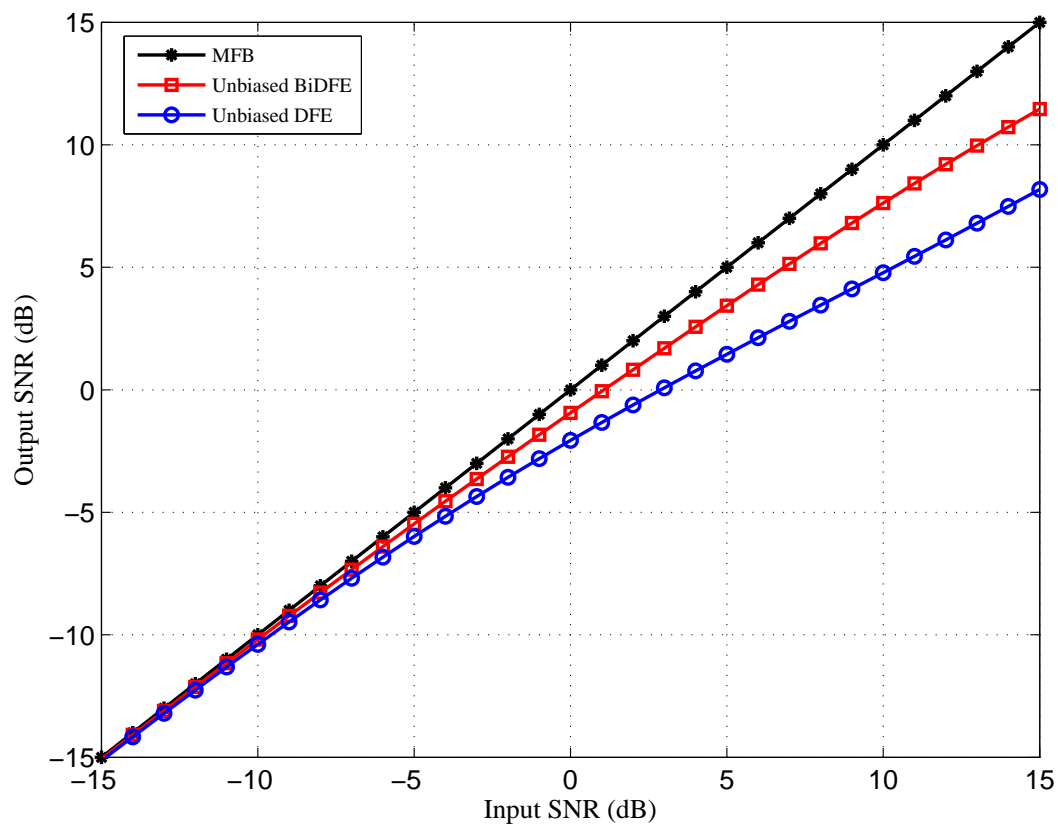Fig. 12: EXIT Chart on the Channel $\mathbf{h_2}$ at a 10 dB with Time-invariant Filters.

Fig. 13: SNR plot on the Channel $\mathbf{h_1}$.